WYDAWNICTWA NAUKOWO-TECHNICZNE
WARSZAWA

Jan Hennel

Podstawy elektroniki półprzewodnikowej

Wydanie 2 poprawione



Redaktor Zuzanna Grzejszczak
Opracowanie graficzne Tadeusz Pietrzyk
Redaktor techniczny Grażyna Miazek

621.382

Podręcznik jest przeznaczony dla studentów wydziałów elektroniki wyższych szkół technicznych. Będzie on również pomocny inżynierom specjalizującym się w zakresie elektroniki półprzewodnikowej.

Tytuł dotowany przez Ministra Edukacji Narodowej

ISBN 83-204-1891-7

WNT Warszawa 1995. Wydanie II
Ark. wyd. 23,3. Ark. druk. 20,5.
Symbol Et/20640/MEN
Drukarnia Wydawnictw Naukowych S.A.
Łódź, ul. Żwirki 2
Zam. 5590

SPIS TREŚCI

PRZEDMOWA 11

WAŻNIEJSZE OZNACZENIA LITEROWE 14

WSTĘP stronica 17

1 DUALIZM FALOWO-KORPUSKULARNY stronica 21

1.1 Elektron w polu elektrycznym/21
1.2 Opis ruchu falowego/24
1.3 Fale elektromagnetyczne/29
1.4 Fotony/31
1.5 Fale de Broglie'a/34
1.6 Zasada nieoznaczoności/38

2 ELEMENTY MECHANIKI KWANTOWEJ stronica 41

2.1 Funkcja falowa/42
2.2 Równanie Schrödingera/43
2.3 Elektron swobodny/44
2.4 Przejście elektronu przez barierę potencjału (przypadek $W < U_0$)/47
2.5* Przejście elektronu przez barierę potencjału (przypadek $W > U_0$)/52
2.6* Zjawisko tunelowe/53

2.7 Elektron w studni potencjału – stany stacjonarne/55
2.8* Elektron w studni potencjału – stany niestacjonarne /60

3 STANY ELEKTRONOWE W ATOMIE stronica 62

3.1 Model atomu Bohra /63
3.2 Kwantowomechaniczna teoria atomu wodoru/66
3.3 Spin elektronu/72
3.4 Atomy wieloelektronowe /73

4 STRUKTURA KRYSTALICZNA CIAŁ STAŁYCH stronica 79

4.1 Ciała krystaliczne/79
4.2 Sieć krystaliczna/ 80
4.3 Układy krystalograficzne/83
4.4 Symbole węzłów, prostych i płaszczyzn sieciowych/84
4.5 Defekty sieci krystalicznej/86
4.6 Drgania sieci krystalicznej/89

5 WIĄZANIA ATOMÓW W KRYSZTALE stronica 92

5.1 Rodzaje wiązań krystalicznych/92
5.2 Wiązanie jonowe/94
5.3 Wiązanie kowalencyjne/97
5.4 Kryształy o wiązaniu kowalencyjnym/100
5.5* Wiązanie spolaryzowane/103
5.6 Metale/107

6 STRUKTURA ENERGETYCZNA CIAŁA STAŁEGO stronica 110

6.1 Rozszczepienie poziomów energetycznych w ciele stałym/110
6.1.1 Metale/110
6.1.2 Półprzewodniki/115
6.2 Elektron w polu periodycznym/120
6.3 Model Kroniga-Penney'a/124
6.4 Ogólne własności pasm energetycznych/130
6.5 Dynamika elektronów w ciele stałym/133

6.6 Pasma paraboliczne/136
6.7 Pojęcie i własności dziury/137

7 PÓŁPRZEWODNIKI stronica 144

7.1 Podstawowe pojęcia i zależności/144
7.2 Rodzaje półprzewodników/149
7.3 Półprzewodniki z prostą przerwą energetyczną/151
7.4 Półprzewodniki ze skośną przerwą energetyczną/157
7.5* Półprzewodniki z zerową przerwą energetyczną/161
7.6 Domieszki w półprzewodnikach/163
7.6.1 Uwagi ogólne/163
7.6.2 Domieszki wodoropodobne/164
7.6.3* Domieszki zlokalizowane/171

8 KONCENTRACJE NOŚNIKÓW ŁADUNKU W PÓŁPRZEWODNIKU W STANIE RÓWNOWAGI TERMODYNAMICZNEJ stronica 174

8.1 Równowaga termodynamiczna/174
8.2 Statystyka Fermiego-Diraca/176
8.3 Koncentracje równowagowe — zależności podstawowe/178
8.4 Prawo działania mas/183
8.5 Półprzewodnik samoistny/184
8.6 Półprzewodniki domieszkowe/188
8.7* Zjawiska występujące przy dużych koncentracjach domieszek/193

9 GENERACJA I REKOMBINACJA NOŚNIKÓW NADMIAROWYCH W PÓŁPRZEWODNIKU stronica 199

9.1 Nośniki nadmiarowe/200
9.2 Rekombinacja bezpośrednia/201
9.3 Rekombinacja pośrednia/208
9.3.1 Pułapki w półprzewodniku/208
9.3.2 Rekombinacja SRH/211
9.3.3* Warunek neutralności elektrycznej/212
9.3.4 Mały poziom zakłócenia/214
9.3.5* Duży poziom zakłócenia/216
9.4 Czas życia nośników nadmiarowych/216
9.5* Rekombinacja Augera/218
9.6* Rekombinacja powierzchniowa/219

10 PRZEPŁYW PRĄDU ELEKTRYCZNEGO W PÓŁPRZEWODNIKU stronica 223

10.1 Podstawowe definicje/223
10.2 Ruchy termiczne nośników w półprzewodniku/225
10.3 Prąd unoszenia/227
10.4 Ruchliwość nośników/235
10.5 Przewodność właściwa/242
10.6 Zjawiska występujące przy dużych natężeniach pola elektrycznego/245
10.6.1* Zależność ruchliwości od natężenia pola elektrycznego/245
10.6.2* Efekt Gunna/247
10.6.3* Jonizacja zderzeniowa/249
10.7 Prąd dyfuzyjny/251

11 ŁADUNEK PRZESTRZENNY W PÓŁPRZEWODNIKU stronica 257

11.1 Równanie Poissona/258
11.2 Półprzewodnik jednorodny/260
11.3 Półprzewodnik niejednorodny/260
11.4 Model pasmowy półprzewodnika niejednorodnie domieszkowanego/264
11.5* Założenie quasi-neutralności elektrycznej/266
11.6 Równania Boltzmanna/269

12 TRANSPORT NOŚNIKÓW NADMIAROWYCH W PÓŁPRZEWODNIKU stronica 272

12.1 Równanie ciągłości/273
12.2 Relaksacja dielektryczna/277
12.2.1 Zakłócenie rozkładu koncentracji nośników większościowych/277
12.2.2 Zakłócenie rozkładu koncentracji nośników mniejszościowych/279
12.3 Ambipolarne równanie transportu/282
12.4 Dyskusja ambipolarnego równania transportu/285
12.5 Przykłady rozwiązania ambipolarnego równania transportu/290
12.5.1 Wstrzykiwanie nośników mniejszościowych/290
12.5.2* Analiza warunku quasi-neutralności/297
12.5.3* Dyfuzja dwuskładnikowa/299

Dodatek A RÓWNANIA MAXWELLA stronica 302
Dodatek B FUNKCJE FALOWE ATOMU WODORU stronica 304
Dodatek C UKŁAD OKRESOWY PIERWIASTKÓW stronica 308
Dodatek D SZEROKOŚĆ PRZERWY ENERGETYCZNEJ PÓŁPRZEWODNIKÓW stronica 310

Dodatek E STOPIEŃ JONIZACJI DOMIESZEK stronica 311
Dodatek F OUASI-POZIOMY FERMIEGO stronica 313
Dodatek G DANE MATERIAŁOWE NAJWAŻNIEJSZYCH PÓŁPRZEWODNIKÓW stronica 315
Dodatek H UZUPEŁNIAJĄCE ZADANIA PROBLEMOWE stronica 316

ROZWIĄZANIA PROBLEMÓW stronica 319
WYKAZ LITERATURY stronica 322
SKOROWIDZ RZECZOWY stronica 324

Żonie mojej Barbarze
pracę tę poświęcam

PRZEDMOWA

Podręcznik ten jest oparty na wykładach „Podstawy elektroniki półprzewodnikowej" prowadzonych przeze mnie od 1980 roku na Wydziale Elektroniki Politechniki Warszawskiej. Wykład ma na celu zapoznanie studentów z podstawowymi własnościami ciał stałych, głównie półprzewodników, w stopniu umożliwiającym słuchanie wykładów z elementów półprzewodnikowych.

Przedmiot „Podstawy elektroniki półprzewodnikowej" ma więc dostarczyć studentom tych wiadomości z fizyki i elektroniki ciała stałego, które są niezbędne do zrozumienia zasad działania takich elementów, jak diody półprzewodnikowe, tranzystory, elementy optoelektroniczne (np. fotodetektory, diody elektroluminescencyjne, baterie słoneczne) i wiele innych. Opiera się on na wiadomościach z fizyki w zakresie szkoły średniej oraz na znajomości rachunku różniczkowego i całkowego w zakresie programu sem. 1 na wydziałach elektroniki wyższych szkół technicznych. W związku z tym wykład posługuje się w możliwie małym stopniu skomplikowanym formalizmem mechaniki kwantowej i obejmuje tylko te zagadnienia z fizyki ciała stałego, które znajdują bezpośrednie zastosowanie w elementarnej teorii elementów półprzewodnikowych.

W podręczniku, podobnie jak w wykładzie, główną uwagę zwrócono na możliwie ścisłe i wyczerpujące, a jednocześnie przystępne omówienie zjawisk elektrycznych i optycznych zachodzących w półprzewodnikach. Do zrozumienia działania tranzystorów i innych elementów półprzewodnikowych konieczna jest znajomość odpowiedzi na następujące 3 pytania:

1) jakie nośniki ładunku i w jakich ilościach występują w poszczególnych obszarach danego elementu,

2) jakie są mechanizmy przemieszczania się tych nośników,

3) jakie prądy powstają w obwodach zewnętrznych elementu wskutek przepływu nośników ładunku przez poszczególne jego obszary.

Zagadnieniom tym poświęcono w książce dużo uwagi. Zasadnicza część wykładu rozpoczyna się od budowy krystalicznej ciała stałego (rozdz. 4), a kończy na analizie równań opisujących transport nośników nadmiarowych w półprzewodniku (rozdz. 12).

Pierwsze 3 rozdziały książki stanowią elementarne wprowadzenie w zagadnienia mechaniki kwantowej, bez której nie można zrozumieć wielu zjawisk zachodzących w ciele stałym. Czytelnik znający równanie Schrödingera i jego najprostsze zastosowania może opuścić te rozdziały.

Zakres podręcznika jest nieco szerszy od wykładu. Dzięki temu może on służyć także studentom wyższych lat studiów jako pomoc w wykładach dotyczących specjalnych rodzajów elementów półprzewodnikowych, nie omawianych w kursie podstawowym. Przykładowo można tu wymienić mikrofalowe elementy półprzewodnikowe (np. diody Gunna) oraz lasery półprzewodnikowe. Podręcznik może być również wykorzystany jako książka pomocnicza (reference book) w przedmiotach związanych z projektowaniem elementów półprzewodnikowych i układów scalonych. Zagadnienia wykraczające poza obowiązujący program wykładu wydrukowano petitem, lub też odpowiedni numer punktu zaznaczono gwiazdką (*). Te fragmenty książki mogą być opuszczone bez szkody dla zrozumienia zasadniczego tekstu.

W podręczniku nie uwzględniono, ze względu na brak miejsca, dwóch ważnych działów fizyki ciała stałego. Dotyczy to obszernej dziedziny zjawisk termoelektrycznych, uwarunkowanych wzajemnym oddziaływaniem procesów cieplnych i elektrycznych (np. zjawiska Seebecka, Peltiera i Thomsona) oraz zjawisk galwanomagnetycznych, wywołanych jednoczesnym działaniem prądu elektrycznego i pola magnetycznego (zjawiska Halla, Ettinghousena, Nernsta, zjawisko magnetooporowe i wiele innych). Zagadnienia te leżą jednakże poza głównym nurtem wykładów prowadzonych na wydziałach elektroniki.

W każdym rozdziale książki zamieszczono pewną liczbę zadań problemowych, z których część jest rozwiązana w tekście, większość zaś — pozostawiona Czytelnikowi do samodzielnego rozwiązania. Taki sposób opracowania podręcznika, w którym wykład jest przeplatany problemami rachunkowymi, ma w zamyśle autora następujące dwie ważne funkcje do spełnienia: po pierwsze, ma wskazywać Czytelnikowi zagadnienia pojęciowo nowe, które powinien On zrozumieć i przyswoić, oraz wnioski, które można wydedukować z poznanych teorii drogą odpowiednich przeliczeń matematycznych, po drugie zaś powinien nakłonić Czytelnika do aktywnego, a nie tylko biernego korzystania z podręcznika. Autor namawia zatem gorąco Czytelnika do samodzielnego rozwiązania wszystkich zawartych w podręczniku zadań problemowych. Prawidłowe odpowiedzi oraz ewentualne komentarze są zamieszczone na końcu książki.

W zakończeniu pragnę podziękować wszystkim, którzy przyczynili się do powstania tej książki oraz przekazali mi swoje uwagi i sugestie. Dr Wiesław Kuźnicz, dr Andrzej Pfitzner i dr Jerzy Woźnicki z Instytutu Technologii Elektronowej Politechniki Warszawskiej zadali sobie trud przeczytania maszynopisu poszczególnych rozdziałów książki. Ich konstruktywny krytycyzm przyczynił się do ulepszenia pierwotnego tekstu. Jestem im za to bardzo wdzięczny.

Szczególnie serdecznie dziękuję mojemu synowi Andrzejowi, pracownikowi Instytutu Fizyki Doświadczalnej Uniwersytetu Warszawskiego, za wiele godzin rozmów i dyskusji nad zagadnieniami będącymi przedmiotem tej książki.

Jestem również bardzo zobowiązany Recenzentom podręcznika: prof. dr. hab. Jerzemu Ginterowi z Uniwersytetu Warszawskiego i prof. dr. hab. Andrzejowi Mulakowi z Politechniki Wrocławskiej za wnikliwe recenzje i cenne uwagi.

AUTOR

WAŻNIEJSZE OZNACZENIA LITEROWE

Alfabet łaciński

A — amplituda, pole powierzchni
a — współczynnik, rozmiar
a_0 — $5{,}292 \cdot 10^{-11}$ m — promień Bohra
B — indukcja magnetyczna
b — współczynnik, rozmiar
C — współczynnik, pojemność
$c = 2{,}998 \cdot 10^{8}$ m/s — prędkość światła w próżni
D — indukcja elektryczna
d — odległość
E — natężenie pola elektrycznego
e — podstawa logarytmów naturalnych
F — siła
f — częstotliwość
G — szybkość generacji nośników
G_{th} — szybkość generacji termicznej
g — szybkość generacji nośników nadmiarowych
H — natężenie pola magnetycznego
$h = 6{,}626 \cdot 10^{-34}$ J/Hz — stała Plancka
$\hbar = h/2\pi = 1{,}0546 \cdot 10^{-34}$ J·s
i — prąd elektryczny
J — gęstość prądu
k — wektor falowy
$k = 1{,}3807 \cdot 10^{-23}$ J/K — stała Boltzmanna
L — długość, droga dyfuzji nośników
l — orbitalna liczba kwantowa
m — masa, magnetyczna liczba kwantowa
$m_0 = 9{,}110 \cdot 10^{-31}$ kg — masa spoczynkowa elektronu
N — liczba, koncentracja domieszek, efektywna gęstość stanów
n — liczba naturalna, główna liczba kwantowa, koncentracja elektronów
P — radialna gęstość prawdopodobieństwa

p – pęd, koncentracja dziur
Q – ładunek
$q = 1{,}602 \cdot 10^{-19}$ C – ładunek elementarny
R – szybkość rekombinacji nośników
r – promień
S – natężenie fali
s – prędkość rekombinacji powierzchniowej
T – temperatura, okres
t – czas
U – energia potencjalna
U_{rg} – wypadkowa szybkość procesów rekombinacyjno-generacyjnych
U_T – potencjał termiczny
u – napięcie
V – potencjał
v – prędkość
W – energia
w – wymiar, gęstość energii
x – współrzędna
y – współrzędna
z – współrzędna

Alfabet grecki

α – kąt, współczynnik
β – kąt, współczynnik
γ – kąt
ε – przenikalność elektryczna
$\varepsilon_0 = 8{,}854 \cdot 10^{-12}$ F/m – przenikalność elektryczna próżni
η – współczynnik
ϑ – kąt
λ – długość fali
μ – ruchliwość nośników, przenikalność magnetyczna
$\mu_0 = 4\pi \cdot 10^{-7}$ H/m – przenikalność magnetyczna próżni
ν – częstotliwość promieniowania elektromagnetycznego
ϱ – gęstość ładunku przestrzennego
σ – przewodność elektryczna (konduktywność)
τ – czas, czas życia nośników
φ – kąt, faza, potencjał
Ψ – funkcja falowa
ψ – amplitudowa funkcja falowa
ω – pulsacja

Indeksy

a – akceptorów
c – pasma przewodnictwa
d – donorów, dyfuzyjny
F – Fermiego
g – przerwy energetycznej
i – samoistny
m – maksymalny
n – elektronów

0 – początkowy, dotyczący stanu równowagi
p – dziur
t – stanów pułapkowych
th – termiczny
u – unoszenia
v – pasma walencyjnego

Znaki matematyczne

exp – symbol funkcji wykładniczej ($\exp x = e^x$)
j – jednostka urojona
ln – logarytm naturalny
Δ – przyrost
∇ – nabla (operator Hamiltona)
$\langle x \rangle$ – wartość średnia zmiennej x

WSTĘP

Książka ta jest poświęcona głównie półprzewodnikom, dlatego też na wstępie należy określić, co rozumiemy przez to pojęcie i jakie są najbardziej istotne cechy tej klasy materiałów.

Spotyka się różne definicje półprzewodników. Według jednej z nich półprzewodniki są to ciała stałe, najczęściej krystaliczne, których przewodność właściwa (konduktywność) w temperaturze pokojowej zawiera się w granicach od 10^{-8} do 10^{4} $(\Omega \cdot cm)^{-1}$, jest więc mniejsza niż przewodność metali (10^{4}–10^{6} $(\Omega \cdot cm)^{-1}$), a większa niż przewodność dielektryków (10^{-20}–10^{-8} $(\Omega \cdot cm)^{-1}$). Najistotniejszą cechą charakterystyczną półprzewodników jest jednakże nie wartość przewodności właściwej, lecz jej znaczna zależność od oddziaływań zewnętrznych.

Prąd elektryczny przepływający przez dany materiał w określonych warunkach zależy od ilości występujących w nim swobodnych nośników ładunku i od prędkości, z jakimi się one poruszają. Prędkości te są na ogół proporcjonalne do natężenia pola elektrycznego, przy czym współczynnik proporcjonalności, nazywany ruchliwością, zależy od rodzaju materiału i od różnych czynników zewnętrznych (np. temperatury). Zakres zmian ruchliwości jest jednakże stosunkowo mały w porównaniu z zakresem, w jakim mogą zawierać się przewodności. Dla określonego materiału ruchliwość może się zmieniać na przykład w granicach od 3 do 4 rzędów wielkości, a dla różnych materiałów w ustalonych warunkach – w granicach od 5 do 6 rzędów wielkości. Przewodności właściwe różnych materiałów znajdujących się w różnych warunkach zewnętrznych zawierają się natomiast w granicach kilkudziesięciu rzędów wielkości (zob. np. rys. 10/13). Tak wielki zakres zmian przewodności wynika głównie z różnych wartości koncentracji nośników ładunku[1]. Podział ciał stałych na metale, półprzewodniki i dielektryki powinien więc uwzględniać tę właśnie wielkość.

---

[1] Koncentracja cząstek danego rodzaju jest to liczba cząstek przypadająca na jednostkę objętości.

W typowych metalach koncentracje swobodnych elektronów zawierają się w granicach od 10^{22} do $2{,}4 \cdot 10^{23}$ cm^{-3} i praktycznie nie zależą od temperatury, oświetlenia i innych czynników zewnętrznych.

Z kolei w dielektrykach koncentrącje swobodnych nośników ładunku są znikomo małe. O właściwościach elektrycznych tych materiałów decydują ładunki związane, które mogą wykonywać tylko ograniczone ruchy w stosunku do swych położeń równowagi. W wyniku tych przemieszczeń dielektryk ulega polaryzacji elektrycznej.

Półprzewodniki natomiast charakteryzują się tym, że koncentracje nośników ładunku są w bardzo silny sposób uzależnione od różnych oddziaływań zewnętrznych. Koncentracje te można zmieniać w szerokich granicach (np. kilkudziesięciu rzędów wielkości), zmieniając temperaturę półprzewodnika lub natężenie padającego nań promieniowania elektromagnetycznego. Nawet ściskanie lub rozciąganie półprzewodnika może w istotny sposób wpływać na koncentracje występujących w nim nośników.

Półprzewodniki są również niezwykle wrażliwe na obecność tzw. domieszek, tzn. obcych atomów, nawet w ilościach, które w innych sytuacjach można uznać za pomijalnie małe. Przykładowo, jeżeli w krysztale krzemu zastąpić co milionowy atom krzemu atomem fosforu, to koncentracja elektronów w temperaturze pokojowej wzrośnie 5 milionów razy.

Biorąc pod uwagę omówione właściwości półprzewodników, można sformułować następującą definicję: *półprzewodniki są to ciała stałe, w których koncentracje swobodnych nośników ładunku mogą osiągać duże wartości i są bardzo silnie zależne od takich czynników, jak temperatura, natężenie promieniowania, koncentracja domieszek itd.*

Termin „półprzewodnik" pojawił się w literaturze naukowej dopiero na początku drugiej dekady bieżącego stulecia. Materiały te były jednak badane, a występujące w nich zjawiska odkrywane wiele dziesiątków lat wcześniej. Już w 1822 roku Seebeck zaobserwował powstawanie siły termoelektrycznej w obwodzie zawierającym siarczek ołowiu lub antymonek cynku. W 1833 r. Faraday stwierdził, że przewodność elektryczna niektórych substancji (np. siarczku srebra) wzrasta przy podwyższaniu temperatury, a w 1839 r. Becquerel odkrył zjawisko fotowoltaiczne w złączu utworzonym przez półprzewodnik zanurzony w elektrolicie. Zjawisko fotowoltaiczne w złączu *p-n* zostało odkryte przez Adamsa i Daya w 1876 r., a trzy lata wcześniej Smith i Sale zaobserwowali fotoprzewodnictwo selenu. Pierwszym wykorzystanym w praktyce elementem półprzewodnikowym był fotodetektor selenowy zastosowany w 1875 r. przez Siemensa w skonstruowanym przez niego fotometrze.

Właściwości prostownicze złącza metal-półprzewodnik zostały odkryte w 1874 r. przez Brauna, który używał do tego celu kryształu siarczku ołowiu i ostrza wolframowego. Minęło jednakże 30 lat zanim detektor krystaliczny, nazywany wówczas „kryształkiem" znalazł swe pierwsze zastosowanie w odbiorniku radiowym.

Fizyka ciała stałego jako ugruntowana nauka rozwinęła się w latach dwudziestych i trzydziestych bieżącego stulecia. Stało się to możliwe dzięki opanowaniu technologii wytwarzania monokryształów oraz metod kontrolowanego wprowadzania do nich określonych domieszek. Drugim niezmiernie ważnym czynnikiem, który zapoczątkował rozwój fizyki ciała stałego, było stworzenie potężnego narzędzia analitycznego, umożliwiającego prawidłową interpretację procesów fizycznych zachodzących w atomach, cząsteczkach i ciałach stałych. Narzędziem tym jest mechanika kwantowa, której podstawy opracowali w latach 1925-1926 Schrödinger, Heisenberg, Dirac i wielu innych badaczy.

Mechanika klasyczna, która panowała niepodzielnie w fizyce do końca XIX wieku, jest całkowicie nieprzydatna do wyjaśnienia zjawisk zachodzących w świecie mikrocząstek (tzn. atomów, elektronów itd.). Można się o tym przekonać, rozpatrując na przykład ruch elektronu w krysztale. Posługując się prawami mechaniki klasycznej, należałoby przyjąć, że elektron poruszający się w krysztale zderza się z każdym atomem napotkanym na swej drodze. W tych warunkach tor elektronu miałby kształt zygzakowaty, a średnia droga swobodna (tzn. średnia odległość między kolejnymi zderzeniami) byłaby rzędu odległości międzyatomowej, a więc rzędu kilkuset pikometrów[1)].

W rzeczywistości średnia droga swobodna elektronu np. w metalu w temperaturze pokojowej jest rzędu kilkudziesięciu nanometrów, a w temperaturach bliskich zera bezwzględnego może wzrosnąć nawet do pojedynczych centymetrów. Jest to wartość wiele milionów razy większa niż odległość między atomami w sieci krystalicznej. Doskonale regularny kryształ nie stawia w ogóle żadnych przeszkód ruchowi elektronu, co oznacza, że takie ciało stałe jest dla elektronu całkowicie przezroczyste.

Zjawiska te wyjaśnia w sposób zgodny z doświadczeniem mechanika kwantowa, w której elektrony i inne mikrocząstki nie są traktowane jako małe kuleczki, lecz jako twory mające zarówno naturę korpuskularną, jak i falową.

[1)] 1 pm = 10^{-12} m. W wielu książkach długości rzędu średnicy atomu lub odległości międzyatomowych są wyrażane w angstremach (1Å = 100 pm).

1

DUALIZM FALOWO-KORPUSKULARNY

Jak już zaznaczono we wstępie, zachowanie się elektronów w ciele stałym może być opisane za pomocą praw mechaniki kwantowej. Jest to dział fizyki zajmujący się badaniem *mikrocząstek*, tj. obiektów o rozmiarach atomu lub mniejszych. Jednym z podstawowych postulatów mechaniki kwantowej jest przypisywanie każdej fali pewnych własności korpuskularnych, a każdej cząstce pewnych własności falowych. Postulaty te zrewolucjonizowały fizykę klasyczną, która rozróżnia dwa zasadniczo różne rodzaje materii, tj. substancję i pole elektromagnetyczne, opisując zachowanie się pierwszej prawami Newtona, drugiej zaś – równaniami Maxwella.

Zanim przystąpimy do przedstawienia niektórych elementów mechaniki kwantowej, niezbędnych do zrozumienia własności półprzewodników, rozważymy w pierwszych 3 punktach tego rozdziału zachowanie się cząstek i fal według opisu fizyki klasycznej. Jako przykład zachowania się cząstki przeanalizujemy ruch swobodnego elektronu w polu elektrycznym, jako przykład fali zaś – falę elektromagnetyczną.

1.1 ELEKTRON W POLU ELEKTRYCZNYM

Elektron jest cząstką elementarną o ujemnym ładunku elektrycznym, równym (co do wartości bezwzględnej) *ładunkowi elementarnemu*

$$q = 1{,}602 \cdot 10^{-19}\ \mathrm{C}$$

i o masie spoczynkowej

$$m_0 = 9{,}110 \cdot 10^{-31}\ \mathrm{kg}$$

W książce tej na ogół nie ma konieczności uwzględniania efektów relatywistycznych. Będziemy tu zatem przyjmowali, że masa swobodnego elektronu wynosi m_0 niezależnie od jego prędkości.

Na elektrony i inne cząstki naładowane (np. jony) działa w polu elektrycznym siła

$$F = QE \tag{1/1}$$

gdzie: Q – ładunek elektryczny danej cząstki (dla elektronu $Q = -q$), $\boldsymbol{E}$ zaś – natężenie pola elektrycznego.

Pod działaniem tej siły cząstka naładowana uzyskuje przyspieszenie skierowane wzdłuż linii pola. Ze względu na ujemny ładunek elektronu, siła wywierana na niego przez pole elektryczne jest skierowana przeciwnie niż wektor natężenia pola. Siła ta stara się więc przesunąć elektron z obszaru o niższym potencjale do obszaru o potencjale wyższym.

W celu wyznaczenia toru i innych parametrów ruchu elektronu w polu elektrycznym, należy, zgodnie z drugą zasadą dynamiki, rozwiązać równanie uzależniające przyspieszenie elektronu od działającej na niego siły. W zapisie wektorowym równanie to, nazywane *równaniem ruchu*, ma następującą postać:

$$m_0 \frac{d\boldsymbol{v}}{dt} = -q\boldsymbol{E} \tag{1/2}$$

Jest ono równoważne układowi trzech równań dla poszczególnych składowych w obranym układzie współrzędnych. Dla współrzędnych prostokątnych x, y, z równania te przedstawiają się następująco:

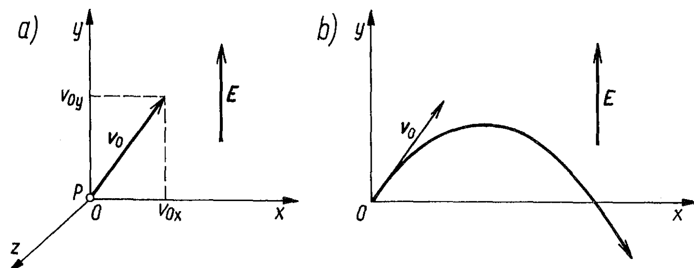
$$\left.\begin{aligned} m_0 \frac{d^2x}{dt^2} &= -qE_x \\ m_0 \frac{d^2y}{dt^2} &= -qE_y \\ m_0 \frac{d^2z}{dt^2} &= -qE_z \end{aligned}\right\} \tag{1/3}$$

Jeżeli jest znany rozkład pola elektrycznego, tzn. są znane wartości składowych E_x, E_y, E_z w każdym punkcie rozpatrywanego obszaru i w każdej chwili

$$\left.\begin{aligned} E_x &= E_x(x, y, z, t) \\ E_y &= E_y(x, y, z, t) \\ E_z &= E_z(x, y, z, t) \end{aligned}\right\} \tag{1/4}$$

to układ równań (1/3) może być rozwiązany dla elektronu o dowolnie zadanych warunkach początkowych, obejmujących jego położenie i prędkość w chwili początkowej. Otrzymane rozwiązanie opisuje w sposób całkowity i jednoznaczny zachowanie się elektronu w rozpatrywanym polu. Jak zobaczymy dalej, mechanika kwantowa wprowadza do tego stwierdzenia pewien stopień nieokreśloności.

Problem 1.1. Jako przykład rozwiązania równań ruchu rozważmy najprostszy przypadek elektronu poruszającego się w niezmiennym w czasie jednorodnym polu elektrycznym o natężeniu $\boldsymbol{E}$. Przyjmijmy, że w chwili $t = 0$ elektron znajduje się w punkcie P i ma prędkość równą $\boldsymbol{v}_0$.



1/1 Ruch elektronu w polu elektrostatycznym: a) wybór układu współrzędnych; b) tor elektronu

Rozwiązanie: Nie tracąc nic na ogólności rozwiązania, przyjmiemy dla wygody, że początek układu współrzędnych x, y, z znajduje się w punkcie P, kierunek osi y – pokrywa się z kierunkiem wektora E, a kierunek osi x jest taki, że wektor prędkości początkowej v_0 leży w płaszczyźnie xy (rys. 1/1a).

W tych warunkach składowe natężenia pola elektrycznego wynoszą odpowiednio: $E_x = 0$, $E_y = E$, $E_z = 0$, składowe zaś prędkości początkowej: v_{0x} i v_{0y}, natomiast $v_{0z} = 0$. Równania ruchu przybiorą zatem postać:

$$\left.\begin{aligned} m_0 \frac{d^2x}{dt^2} &= 0 \\ m_0 \frac{d^2y}{dt^2} &= -qE \\ m_0 \frac{d^2z}{dt^2} &= 0 \end{aligned}\right\} \qquad (1/5)$$

Dwukrotne całkowanie tych równań przy uwzględnieniu warunków, że w chwili $t = 0$

$$x = y = z = 0$$

oraz

$$\frac{dx}{dt} = v_{0x}, \qquad \frac{dy}{dt} = v_{0y}, \qquad \frac{dz}{dt} = 0$$

prowadzi do rozwiązania o postaci

$$\left.\begin{aligned} x(t) &= v_{0x}\, t \\ y(t) &= \frac{-q}{2m_0} Et^2 + v_{0y}\, t \\ z(t) &= 0 \end{aligned}\right\} \qquad (1/6)$$

Ruch elektronu odbywa się więc w płaszczyźnie xy, a jego torem jest parabola, której oś pokrywa się z kierunkiem pola elektrycznego (rys. 1/1b). Można to łatwo wykazać, eliminując czas t z równań (1/6).

Na podanym przykładzie można stwierdzić, że w fizyce klasycznej ruch elektronu w polu sił jest jednoznacznie wyznaczony (zdeterminowany) przez charakterystykę sytuacji, w jakiej się on odbywa. W rozpatrywanym przykładzie na charakterystykę tę składa się opis rozkładu pola elektrycznego oraz dane dotyczące położenia początkowego i prędkości początkowej elektronu. Zgodnie z koncepcją determinizmu, w tych samych warunkach zachodzą te same zjawiska, a zatem każdy elektron będzie się zachowywał identycznie.

Energia potencjalna cząstki naładowanej znajdującej się w polu elektrostatycznym wyraża się wzorem

$$W_p = QV \qquad (1/7)$$

gdzie V – potencjał elektrostatyczny.

Jak wiadomo, zarówno potencjał, jak i energia potencjalna są określone z dokładnością do stałej addytywnej. Oznacza to, że punkt odniesienia, względem którego mierzone są te wielkości, może być obrany w sposób dowolny. Często przyjmuje się, że potencjał w nieskończoności jest równy zeru. W książce tej spotkamy się jednak również z innymi sposobami ustalania punktu odniesienia potencjału. Czynność tę nazywamy *normowaniem* potencjału.

Dla elektronu wzór (1/7) przyjmuje postać

$$\boxed{W_p = -qV} \qquad (1/8)$$

Ze względu na ścisły związek zachodzący między energią potencjalną elektronu a potencjałem, dogodnie jest wyrażać energie elektronów (oraz innych cząstek, np. fotonów) w *elektronowoltach*. Elektronowolt jest równy energii nabytej przez elektron na drodze, na której końcach występuje różnica potencjałów równa 1 V. Znając ładunek elektronu, można napisać, że

$$1 \text{ eV} = 1{,}602 \cdot 10^{-19} \text{ J}$$

1.2 OPIS RUCHU FALOWEGO

Fala jest to zaburzenie pola fizycznego (np. elektromagnetycznego) rozchodzące się ze skończoną prędkością i przenoszące energię. Matematycznie falę opisuje funkcja uzależniająca zmiany określonej wielkości fizycznej A od położenia i czasu. Funkcja ta jest rozwiązaniem tzw. *równania falowego*[1].

Wielkość A może być skalarem (w przypadku fal akustycznych jest to np. ciśnienie) lub wektorem (np. natężenie pola elektrycznego w przypadku fal elektromagnetycznych).

Przy opisie wszelkich fal podstawowe znaczenie ma pojęcie *fali płaskiej*, w której wielkość A ma jednakową wartość w każdym punkcie płaszczyzny prostopadłej do pewnego wektora jednostkowego $\boldsymbol{n}$, określającego kierunek rozchodzenia się fali. W ośrodku nie wykazującym dyspersji[2] fala ta może być opisana wyrażeniem

$$A(\boldsymbol{r}, t) = F(\boldsymbol{n} \cdot \boldsymbol{r} - vt) \qquad (1/9)$$

gdzie: F – dowolna funkcja, której druga pochodna jest ciągła; $\boldsymbol{r}$ – wektor określający położenie rozpatrywanego punktu w stosunku do ustalonego punktu odniesienia (tzw. *wektor wodzący*); v – prędkość rozchodzenia się fali.

---

[1] Dla fal elektromagnetycznych równanie to podano w dodatku A.

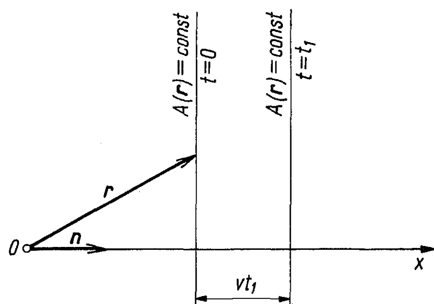
[2] Zob. odnośnik na s. 120.

Zmiany rozpatrywanej wielkości fizycznej mogą zachodzić w kierunku rozchodzenia się fali (fale takie nazywamy *podłużnymi*), albo też odbywać się w płaszczyźnie prostopadłej do kierunku rozchodzenia się fali (fale *poprzeczne*). Przykładem fal podłużnych są fale akustyczne, poprzecznych zaś – fale elektromagnetyczne w nieograniczonym idealnym dielektryku.

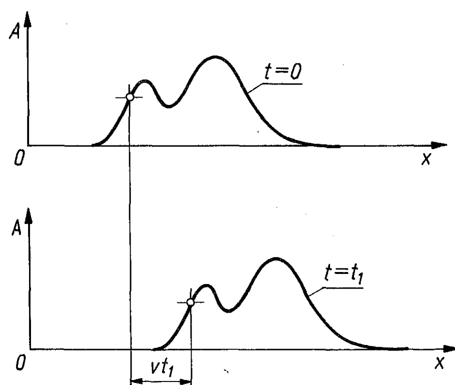
Przyjmując początek wektora wodzącego za początek układu współrzędnych, a kierunek wektora $\boldsymbol{n}$ za kierunek osi x (rys. 1/2), można zauważyć, że iloczyn skalarny $\boldsymbol{n}\cdot\boldsymbol{r}$ jest równy współrzędnej x danego punktu. Wyrażenie (1/9) przyjmuje zatem postać

$$A = F(x-vt) \qquad (1/10)$$

z której wynika, że płaszczyzna stałej wartości A (nazywana *płaszczyzną falową*) przesuwa się wzdłuż osi x z prędkością v. Przykład zależności A (x) dla 2 wartości czasu t przedstawiono na rys. 1/3.



1/2 Fala płaska rozchodząca się w kierunku osi x



1/3 Przebieg A (x) dla 2 wartości czasu

Bardzo ważnym rodzajem fali jest fala *harmoniczna* (nazywana również falą *monochromatyczną*), w której rozpatrywana wielkość fizyczna wykonuje w każdym punkcie przestrzeni drgania sinusoidalne. Odgrywa ona podstawową rolę, ponieważ każdy przebieg okresowy można traktować jako wynik nakładania się skończonej lub nieskończonej ilości przebiegów harmonicznych.

Płaska fala harmoniczna rozchodząca się wzdłuż osi x może być opisana równaniem

$$A = A_m \cos\left[k(x-vt)+\varphi_0\right] \qquad (1/11)$$

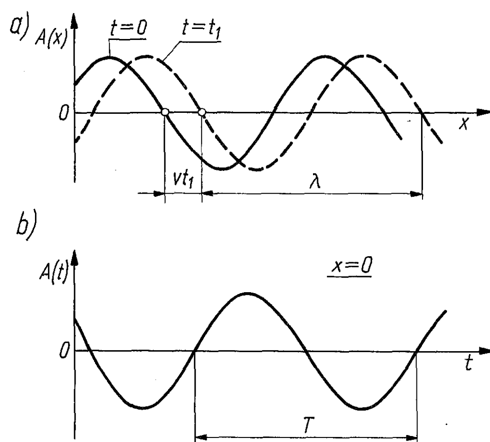
gdzie: A_m oznacza *amplitudę* fali, a argument funkcji cosinus

$$\varphi = k(x-vt)+\varphi_0 \qquad (1/12)$$

fazę fali. Wielkość φ_0 jest fazą początkową.

Przebieg zależności A (x) przy t = const oraz A (t) przy x = const dla fali opisanej wzorem (1/11) przedstawiono na rys. 1/4.

Ze wzoru (1/12) wynika, że faza fali zmienia się jednostajnie zarówno w czasie, jak i w przestrzeni (w funkcji współrzędnej x). Wielkość charakteryzu-



1/4 Płaska fala harmoniczna: a) przebieg $A(x)$ dla 2 wartości czasu; b) przebieg $A(t)$ dla $x = \text{const}$

jącą szybkość zmian fazy fali w przestrzeni, tzn. pochodną

$$k = \left|\frac{\partial \varphi}{\partial x}\right| \tag{1/13}$$

nazywamy *liczbą falową* (lub *współczynnikiem fazowym*) fali.

Faza fali zmienia się o kąt 2π na odcinku o długości λ, którą nazywamy *długością fali*, a zatem

$$k = \frac{2\pi}{\lambda} \tag{1/14}$$

Wektorem falowym nazywamy wektor $\boldsymbol{k}$, którego moduł jest określony wzorem (1/14), kierunek zaś jest zgodny z kierunkiem rozchodzenia się fali, tzn.

$$\boldsymbol{k} = k\boldsymbol{n} = \frac{2\pi}{\lambda}\boldsymbol{n} \tag{1/15}$$

Szybkość zmian fazy w czasie, tzn. pochodną

$$\omega = \left|\frac{\partial \varphi}{\partial t}\right| \tag{1/16}$$

nazywamy *pulsacją* fali. W czasie równym okresowi drgań T faza fali zmienia się o kąt 2π, a zatem

$$\omega = \frac{2\pi}{T} = 2\pi f \tag{1/17}$$

gdzie $f = 1/T$ oznacza *częstotliwość* fali.

Na podstawie wzoru (1/12) można otrzymać następujące wyrażenie na prędkość rozchodzenia się fali:

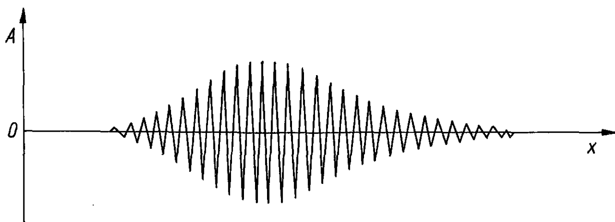
$$v = \frac{\omega}{k} \tag{1/18}$$

Dla fali harmonicznej jest to prędkość, z jaką ustalona wartość fazy przesuwa się wzdłuż kierunku rozchodzenia się fali. Dlatego też nosi ona nazwę *prędkości fazowej*.

Podstawiając (1/18) do wzoru (1/11), otrzymamy następujące częściej stosowane wyrażenie na płaską falę harmoniczną:

$$A = A_m \cos(kx - \omega t + \varphi_0) \tag{1/19}$$

Idealna fala harmoniczna jest tworem nierealnym, jest ona bowiem nieograniczona zarówno w przestrzeni, jak i w czasie. Rzeczywiste fale występujące w przyrodzie są oczywiście ograniczone do pewnego obszaru przestrzennego i do pewnego przedziału czasowego. W związku z tym nie mogą one być idealnie monochromatyczne, tzn. nie mogą być opisane funkcją o postaci (1/19) o ściśle określonej (tzw. *dyskretnej*) częstotliwości. Falę taką można uważać za przebieg będący superpozycją nieskończenie wielu fal harmonicznych o częstotliwościach zawartych w pewnym obszarze widma. Obszar ten jest tym szerszy, im krótszy jest dany ciąg falowy. Jeżeli szerokość widma częstotliwości jest niewielka w stosunku do częstotliwości środkowej, to mówimy o *grupie* fal harmonicznych, czyli o *paczce falowej*. Przykład takiej paczki pokazano na rys. 1/5.



1/5 Paczka falowa

W celu zbadania własności propagacyjnych takiej paczki, rozważmy najprostszy przypadek, gdy badany przebieg jest superpozycją dwóch fal sinusoidalnych o jednakowych amplitudach i zerowych fazach początkowych, natomiast nieznacznie różniących się pulsacjach $\omega_0 - \Delta\omega$ i $\omega_0 + \Delta\omega$ oraz liczbach falowych $k_0 - \Delta k$ i $k_0 + \Delta k$.

Fala wypadkowa będzie miała przebieg opisany następującym wzorem:

$$A = A_m \cos[(k_0 - \Delta k)\,x - (\omega_0 - \Delta\omega)\,t] + \\ + A_m \cos[(k_0 + \Delta k)\,x - (\omega_0 + \Delta\omega)\,t] \tag{1/20}$$

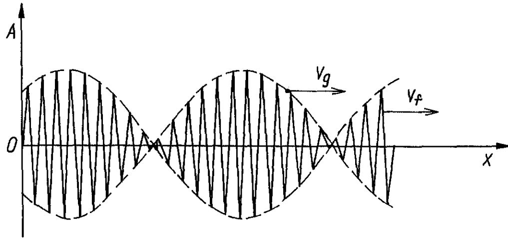
który, po przeprowadzeniu odpowiednich przekształceń trygonometrycznych, może być doprowadzony do postaci

$$A = 2A_m \cos(\Delta k\, x - \Delta\omega\, t) \cos(k_0\, x - \omega_0\, t) \tag{1/21}$$

Z otrzymanego wzoru wynika, że falę wypadkową można traktować w przybliżeniu jako falę o liczbie falowej k_0 i pulsacji ω_0, z tym, że amplituda fali zmienia się proporcjonalnie do wartości funkcji $|\cos(\Delta k\, x - \Delta\omega\, t)|$ — rys. 1/6.

Faza fali przesuwa się wzdłuż osi x z prędkością fazową

$$v_f = \frac{\omega_0}{k_0} \tag{1/22}$$



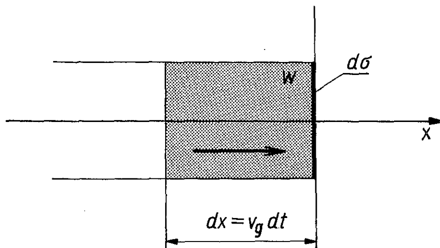
1/6 Fala utworzona przez superpozycję 2 fal harmonicznych o nieznacznie różniących się częstotliwościach
v_f — prędkość fazowa; v_g — prędkość grupowa

jej obwiednia natomiast (linia przerywana na rys. 1/6) z prędkością v_g, zwaną *prędkością grupową*. Jak łatwo stwierdzić na podstawie wyrażenia (1/21)

$$v_g = \frac{\Delta\omega}{\Delta k} \tag{1/23}$$

Prędkość grupowa ma ważny sens fizyczny, jest to bowiem prędkość, z jaką są przenoszone przez falę wszelkie sygnały i zaburzenia. W ośrodkach bezstratnych jest to również prędkość, z jaką odbywa się przenoszenie energii przez falę.

Rozpatrzmy dokładniej przepływ strumienia energii przez płaszczyznę prostopadłą do kierunku rozchodzenia się grupy płaskich fal harmonicznych (rys. 1/7). Weźmy pod uwagę walec o polu podstawy dσ i osi równoległej do kierunku rozchodzenia się fali.



1/7
Przepływ strumienia energii przez płaszczyznę prostopadłą do kierunku rozchodzenia się fali

Niech w oznacza gęstość energii fali (tzn. energię przypadającą na jednostkę objętości) w rozpatrywanym punkcie. Ponieważ paczka falowa przesuwa się wzdłuż osi x z prędkością grupową v_g, zatem w czasie dt energia dW przenoszona przez falę przez powierzchnię dσ wyniesie

$$\mathrm{d}W = w\,\mathrm{d}\sigma\,\mathrm{d}x = w\,v_g\,\mathrm{d}\sigma\,\mathrm{d}t \tag{1/24}$$

Natężeniem fali S nazywamy gęstość strumienia energii przechodzącej przez rozpatrywaną płaszczyznę falową $x = \text{const}$. Jest to więc energia przypadająca na jednostkę czasu i jednostkę powierzchni prostopadłej do kierunku rozchodzenia się fali. Ze wzoru (1/24) wynika następujący ważny związek:

$$S = wv_g \tag{1/25}$$

Występująca w tym wzorze gęstość energii fali w zależy od kwadratu jej amplitudy i od własności ośrodka, w którym rozchodzi się fala.

1.3 FALE ELEKTROMAGNETYCZNE

Falą elektromagnetyczną nazywamy zjawisko rozchodzenia się zaburzeń pola elektromagnetycznego. Pole to jest opisane za pomocą wektora elektrycznego $\boldsymbol{E}$ (natężenia pola elektrycznego) i wektora magnetycznego $\boldsymbol{H}$ (natężenia pola magnetycznego). Podstawowymi równaniami opisującymi zachowanie się wektorów $\boldsymbol{E}$ i $\boldsymbol{H}$ w czasie i w przestrzeni są *równania Maxwella*[1]. Zasadniczym wnioskiem wypływającym z tych równań jest istnienie fal elektromagnetycznych, które w jednorodnym, izotropowym i bezstratnym dielektryku rozchodzą się z prędkością

$$v = \frac{1}{\sqrt{\varepsilon\mu}} \tag{1/26}$$

gdzie: ε – przenikalność elektryczna; μ – przenikalność magnetyczna dielektryka.

Dla próżni przenikalności te, oznaczane symbolami ε_0 i μ_0, wynoszą odpowiednio:

$$\varepsilon_0 = 8{,}854 \cdot 10^{-12}\ \mathrm{F/m}$$

$$\mu_0 = 4\pi \cdot 10^{-7}\ \mathrm{H/m}$$

a prędkość rozchodzenia się fali

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} = 2{,}998 \cdot 10^8\ \mathrm{m/s} \tag{1/27}$$

W polu elektromagnetycznym jest magazynowana energia. Jest ona pobierana ze źródła w czasie wytwarzania pola, a następnie może być wykorzystana np. do pracy związanej z przesuwaniem ładunków elektrycznych znajdujących się w obszarze działania sił pola.

Z równań Maxwella wypływa wniosek, że gęstość energii zmagazynowanej w polu elektromagnetycznym

$$w = \frac{1}{2}(\varepsilon E^2 + \mu H^2) \tag{1/28}$$

gdzie E i H oznaczają odpowiednio chwilowe wartości wektora elektrycznego i magnetycznego w rozpatrywanym punkcie pola.

Uśredniając tę energię dla fali harmonicznej i korzystając z zależności $H = \sqrt{\frac{\varepsilon}{\mu}}\, E$, otrzymuje się następujące wyrażenie na średnią gęstość energii:

$$w_{\text{śr}} = \frac{1}{2}\varepsilon E_m^2 \tag{1/29}$$

Zgodnie z uwagą na końcu p. 1.2, gęstość energii zmagazynowanej w polu elektromagnetycznym jest więc proporcjonalna do kwadratu amplitudy drgań elektromagnetycznych.

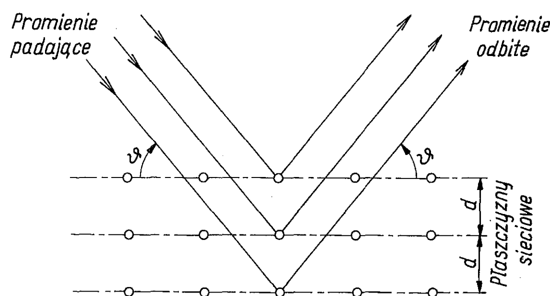
Istnieje wiele zjawisk, w których uwidocznia się falowa natura promieniowania elektromagnetycznego. Należą do nich np. dyfrakcja i interferencja.

[1] Równania te podano w dodatku A.

Dyfrakcja jest to zjawisko związane z odstępstwami rozchodzenia się fali od praw optyki geometrycznej. Ujawnia się ono np. w postaci ugięcia fali na krawędzi przeszkody (lub otworu) umieszczonej na jej drodze. Dyfrakcja występuje tym wyraźniej, im bardziej rozmiary przeszkody (otworu) są zbliżone do długości fali.

Interferencja jest to nakładanie się fal o jednakowych częstotliwościach, prowadzące do wzmocnienia (tzn. zwiększenia amplitudy) lub osłabienia (zmniejszenia amplitudy) fali wypadkowej w danym punkcie w zależności od różnic faz fal składowych. W przypadku 2 interferujących ze sobą fal amplituda fali wypadkowej osiąga wartość maksymalną, równą sumie amplitud fal składowych, w punktach, w których różnica ich faz wynosi $2\pi n$, gdzie n jest liczbą całkowitą. Osiąga natomiast wartość minimalną (równą różnicy amplitud fal składowych) w punktach, w których różnica faz wynosi $(2n+1)\pi$.

Zjawiska dyfrakcji i interferencji mogą być obserwowane np. w siatce dyfrakcyjnej, mającej liczne zastosowania w optyce. Naturalną siatkę dyfrakcyjną dla promieniowania o bardzo małej długości fali (np. promieniowania rentgenowskiego) stanowią kryształy. Jest to siatka przestrzenna, w której obraz dyfrakcyjny dogodnie jest obserwować w promieniowaniu odbitym (rys. 1/8).



1/8
Dyfrakcja promieni rentgenowskich w krysztale

Wiązka promieni padających, ulegając dyfrakcji zostaje częściowo odbita przez płaszczyzny sieciowe kryształu. Wzmocnienie interferencyjne promieni odbitych od kolejnych płaszczyzn zachodzi wówczas, gdy różnice dróg tych promieni są równe całkowitej wielokrotności długości fali w krysztale. Na podstawie rys. 1/8 można stwierdzić, że zachodzi to wówczas, gdy jest spełniony tzw. *warunek Bragga* wyrażony wzorem

$$2d \sin \vartheta = n\lambda \tag{1/30}$$

gdzie: n – liczba całkowita; d – odległość między płaszczyznami sieciowymi kryształu.

Ponieważ $\sin \vartheta$ nie może być większy od jedności, n natomiast mniejsze od jedności, zatem odbicie Bragga nie zachodzi dla fal o długości większej niż $2d$.

Z warunku Bragga będziemy korzystali w rozdz. 6 przy wyjaśnianiu przyczyn powstawania pasm energetycznych w półprzewodnikach.

1.4 FOTONY

Opisane w poprzednim punkcie zjawiska dyfrakcji i interferencji, jak również olbrzymią liczbę innych zjawisk, można wytłumaczyć i dokładnie obliczyć ich wyniki na podstawie elektrodynamiki klasycznej, opartej na równaniach Maxwella. Istnieją jednakże zjawiska, których interpretacja oparta na prawach fizyki klasycznej daje wyniki niezgodne z doświadczeniem, a niekiedy wręcz niedorzeczne.

Tak na przykład stosowanie fizyki klasycznej prowadzi do prawa promieniowania ciała doskonale czarnego (zwanego prawem Rayleigha-Jeansa), zgodnie z którym widmowa zdolność emisyjna tego ciała rośnie monotonicznie ze wzrostem częstotliwości, a wskutek tego całkowita zdolność emisyjna dąży do nieskończoności. Ten absurdalny wniosek, sprzeczny z zasadą zachowania energii, jest określany nazwą katastrofy w nadfiolecie.

W celu pokonania trudności powstających przy interpretacji zjawiska promieniowania temperaturowego, Planck odszedł od fizyki klasycznej, robiąc radykalne, klasycznie nieuzasadnione założenie, że energia układów fizycznych nie może zmieniać się w sposób ciągły, lecz zmienia się skokowo. Zgodnie z postulatem Plancka układ fizyczny o 1 stopniu swobody wykonujący drgania harmoniczne o częstotliwości ν może mieć tylko takie wartości całkowitej energii, które spełniają zależność

$$W = nh\nu \qquad (1/31)$$

gdzie: n – liczba całkowita nieujemna; h – nowa podstawowa stała przyrody, zwana *stałą Plancka*, której wartość

$$h = 6{,}626 \cdot 10^{-34}\ \mathrm{J/Hz}$$

Oparte na powyższej hipotezie prawo promieniowania Plancka opisuje ściśle obserwowany rozkład promieniowania w całym zakresie częstotliwości[1].

Późniejszy rozwój mechaniki kwantowej doprowadził do następującego bardziej poprawnego wzoru na energię oscylatora harmonicznego:

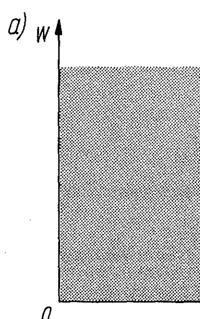
$$W = \left(n + \frac{1}{2}\right) h\nu \qquad (1/32)$$

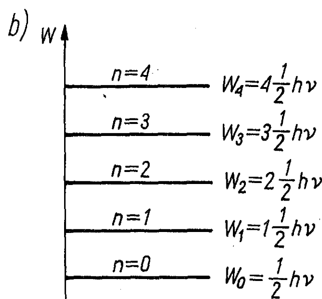
Oznacza to, że nawet w najniższym stanie energetycznym ($n = 0$) oscylator ten ma różną od zera energię $W_0 = \frac{1}{2} h\nu$. Poprawka ta nie ma jednakże istotnego wpływu na prawo promieniowania Plancka.

Wykresy dozwolonych wartości energii oscylatora harmonicznego wg fizyki klasycznej i fizyki kwantowej zestawiono na rys. 1/9. W pierwszym przypadku każda nieujemna wartość energii jest dozwolona – mówimy wówczas, że energia ma *widmo ciągłe*. W drugim przypadku mówimy, że energia jest *skwantowana*, a najmniejszą wartość, o jaką może się ona zmieniać, nazywamy *kwantem energii*.

---

[1] Dzień 14 grudnia 1900 r., w którym Max Planck ogłosił swoje prawo, jest uważany za datę narodzin fizyki kwantowej.

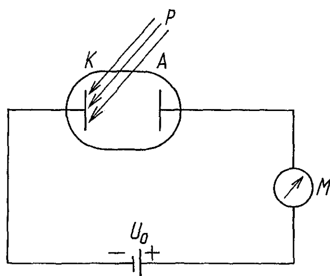




1/9
Widmo energetyczne oscylatora harmonicznego: a) wg fizyki klasycznej; b) wg fizyki kwantowej

Dla obiektów makroskopowych dozwolone wartości energii są rozmieszczone tak gęsto, że nieciągłości energii nie dostrzegamy, ani też nie możemy jej stwierdzić, stosując nawet najczulsze przyrządy. Tylko w układach, których energia jest bardzo mała, a częstotliwość drgań duża, daje się zaobserwować skwantowanie energii.

Innym zjawiskiem, w którym prawa fizyki klasycznej prowadzą do wniosków sprzecznych z doświadczeniem, jest *zjawisko fotoelektryczne*. Polega ono na emisji elektronów pod wpływem promieniowania elektromagnetycznego. Przyrząd do obserwacji tego zjawiska przedstawiono na rys. 1/10. Składa się on z umieszczonych w próżni 2 metalowych elektrod, do których jest doprowadzone stałe napięcie U_0. Elektrodę połączoną z ujemnym biegunem napięcia nazywamy katodą, z dodatnim zaś – anodą. Umieszczony w obwodzie miernik M służy do pomiaru prądu płynącego od anody do katody.



1/10
Przyrząd do badania zjawiska fotoelektrycznego
K — katoda; *A* — anoda; *P* — promieniowanie elektromagnetyczne

Jeżeli katoda nie jest nagrzana do dostatecznie wysokiej temperatury i nie pada na nią promieniowanie elektromagnetyczne, to prąd w obwodzie nie występuje. Pod wpływem natomiast padającego na katodę promieniowania elektromagnetycznego mogą być z niej emitowane elektrony, które wówczas wytwarzają w obwodzie odpowiedni prąd. Aby wydostać się z katody, elektrony muszą pokonać działanie sił wiążących je z katodą, tzn. muszą wykonać pewną pracę zwaną *pracą wyjścia*.

Zgodnie z elektrodynamiką klasyczną elektrony w metalu ulegają oddziaływaniu z falą elektromagnetyczną, której natężenie jest, jak to wynika ze wzorów (1/25) i (1/29), proporcjonalne do E_m^2. Przy zwiększaniu intensywności promieniowania emisja elektronów powinna zatem wystąpić dopiero wówczas, gdy

natężenie fali przekroczy pewną wartość, przy której elektron uzyskuje energię równą pracy wyjścia.

Wniosek ten jest niezgodny z rzeczywistością, ponieważ emisja elektronów występuje przy dowolnie małym natężeniu promieniowania. W sposób natomiast zupełnie niezrozumiały na gruncie fizyki klasycznej jest ona uwarunkowana częstotliwością padającego promieniowania. Występuje mianowicie dopiero wówczas, gdy częstotliwość ta przekroczy pewną wartość progową ν_0, zależną od pracy wyjścia materiału katody.

W celu wytłumaczenia zjawiska fotoelektrycznego, Einstein wysunął hipotezę, że promieniowanie elektromagnetyczne ma w pewnych sytuacjach własności korpuskularne, tzn. że można je traktować jako strumień poruszających się z prędkością światła cząstek (zwanych *fotonami*), z których każda niesie pojedynczy kwant energii.

Dla promieniowania monochromatycznego o częstotliwości ν każdy foton ma energię wyrażoną wzorem

$$\boxed{W = h\nu} \qquad (1/33)$$

Zgodnie z postulatem Einsteina cała energia fotonu zaabsorbowanego w materiale katody zostaje przekazana jednemu elektronowi. Elektron ten może wyswobodzić się z katody tylko wówczas, gdy energia ta jest większa niż praca wyjścia, a to zachodzi wówczas gdy częstotliwość promieniowania przekroczy wartość progową ν_0.

Własności korpuskularne promieniowania elektromagnetycznego ujawniają się także w *reakcjach fotochemicznych*, tzn. reakcjach zachodzących pod wpływem światła, promieniowania ultrafioletowego lub rentgenowskiego. Dobrze znanym przykładem reakcji fotochemicznej jest rozkład halogenków srebra wykorzystywany w procesie fotograficznym. W reakcjach fotochemicznych obowiązuje ustalone przez Einsteina prawo głoszące, że liczba cząsteczek ulegających przemianie fotochemicznej jest równa liczbie fotonów pochłoniętych przez daną substancję.

Problem 1.2. Tak zwana żółta linia widmowa sodu (wykorzystywana, np. w lampach służących do oświetlenia ulic) odpowiada przejściu atomu sodu ze stanu $3p$ do stanu $3s$, przy czym różnica energii tych stanów wynosi 2,1 eV, a czas wypromieniowania fotonu $\tau = 1{,}6\cdot 10^{-8}$ s. Oblicz długość fali tego fotonu oraz długość reprezentującej go paczki falowej.

Rozwiązanie: Energia fotonu wyrażona w dżulach wynosi

$$W = 1{,}6\cdot 10^{-19}\cdot 2{,}1 = 3{,}36\cdot 10^{-19}\ \text{J}$$

Korzystając ze wzoru (1/33), znajdujemy

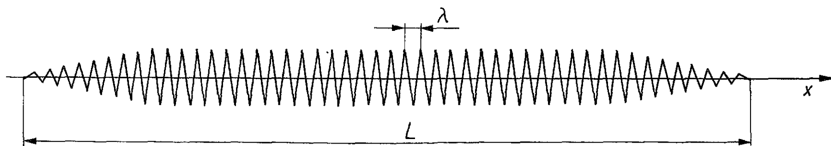
$$\nu = \frac{W}{h} = \frac{3{,}36\cdot 10^{-19}}{6{,}626\cdot 10^{-34}} = 5{,}07\cdot 10^{14}\ \text{Hz}$$

a zatem

$$\lambda = \frac{c}{\nu} = \frac{3\cdot 10^{8}}{5{,}07\cdot 10^{14}} = 5{,}9\cdot 10^{-7}\ \text{m} = 590\ \text{nm}$$

Foton porusza się z prędkością światła c. Skoro czas wypromieniowania fotonu wynosi τ, zatem długość paczki falowej wynosi

$$L = c\tau = 3\cdot 10^{8}\cdot 1{,}6\cdot 10^{-8} = 4{,}8\ \text{m}$$



1/11 Foton jako paczka falowa

Foton można więc wyobrażać sobie w postaci paczki przedstawionej schematycznie na rys. 1/11. Paczka ta ma długość równą 8 milionom długości fali promieniowania.

Problem 1.3. Najsłabsze gwiazdy widoczne nieuzbrojonym okiem wysyłają światło, którego natężenie na powierzchni ziemi jest rzędu $2 \cdot 10^{-9}$ lx. Przyjmując dla uproszczenia, że jest to światło monochromatyczne o barwie żółtej, której odpowiada długość fali 580 nm, oszacuj ile fotonów wpada w tych warunkach do oka w czasie 1/50 s (jest to orientacyjny czas, w którym w oku kumulują się wrażenia wzrokowe). Średnica źrenicy oka adaptowanego do ciemności wynosi ok. 8 mm, a skuteczność świetlna promieniowania o długości fali 580 nm ma wartość 590 lm/W.

Foton, czyli kwant energii promieniowania elektromagnetycznego, jest cząstką elementarną, która ma nie tylko określoną energię, lecz również masę i pęd. Wielkości te można wyznaczyć z równania Einsteina

$$W = mc^2 = pc \tag{1/34}$$

Pęd fotonu wynosi więc

$$p = \frac{h\nu}{c} = \frac{h}{\lambda} \tag{1/35}$$

Pęd fotonu ujawnia się szczególnie wyraźnie w *zjawisku Comptona*, polegającym na rozpraszaniu promieniowania elektromagnetycznego na swobodnych elektronach. Zachodzi ono wskutek zderzenia pojedynczego fotonu z elektronem, przy czym foton przekazuje elektronowi część pędu i energii. W wyniku zderzenia następuje zmiana długości fali promieniowania. Zjawiska tego nie można wyjaśnić na podstawie klasycznej teorii światła, jest ono natomiast jednym z głównych dowodów teorii kwantowej[1)].

Na podstawie zjawisk omawianych w tym i w poprzednim punkcie można stwierdzić, że promieniowanie elektromagnetyczne ujawnia w pewnych sytuacjach naturę falową, w innych zaś – korpuskularną. Własność tę nazywamy *dualizmem falowo-korpuskularnym.*

1.5 FALE DE BROGLIE'A[2)]

Hipoteza fotonowej natury promieniowania elektromagnetycznego była pierwszym wyłomem w fizyce klasycznej i pierwszym krokiem, który prowadził do mechaniki kwantowej. Następnym i decydującym krokiem była wysunięta przez

---

[1)] Na gruncie fizyki klasycznej powinniśmy oczekiwać, że elektron będzie oscylował z częstotliwością padającej fali, a wskutek tego będzie wypromieniowywał falę o tej samej częstotliwości.

[2)] Czytaj: de broja.

de Broglie'a w 1924 r. hipoteza, według której dualizm falowo-korpuskularny dotyczy nie tylko promieniowania, ale także wszelkich cząstek materialnych, a więc elektronów, protonów, atomów itd. Zgodnie z tą hipotezą, każdej cząstce o całkowitej energii W i pędzie $\boldsymbol{p}$ należy przyporządkować falę o częstotliwości ν i wektorze falowym $\boldsymbol{k}$, które spełniają następujące zależności:

$$h\nu = W \tag{1/36}$$

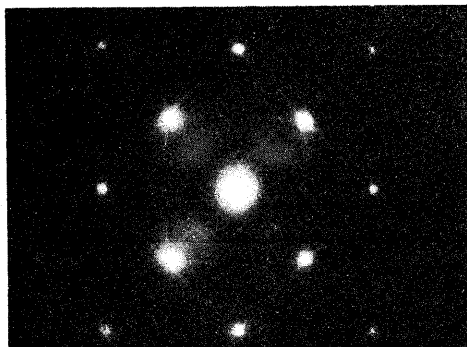
$$\hbar \boldsymbol{k} = \boldsymbol{p} \tag{1/37}$$

gdzie $\hbar$ (h – kreślone) – stała Plancka podzielona przez 2π.

Fale te nazywamy *falami materii* albo falami de Broglie'a. Natura tych fal była początkowo nieznana, ponieważ hipoteza de Broglie'a nie opierała się na żadnych faktach doświadczalnych.

Z ostatniego wzoru wynika, że długość fali przypisywanej danej cząstce jest tym mniejsza, im większy jest pęd tej cząstki, tzn. im większe są jej masa i prędkość, z jaką się ona porusza.

Istnienie fal materii zostało potwierdzone w 1927 r. przez Davissona i Germera, którzy wykazali, że wiązka elektronów odbita od powierzchni kryształu tworzy obraz dyfrakcyjny (rys. 1/12) podobny do obrazu, jaki uzyskuje się za pomocą promieni rentgenowskich. Wkrótce potem Estermann i Stern stwierdzili, że również atomy helu i cząsteczki wodoru ulegają dyfrakcji zgodnie z hipotezą de Broglie'a.



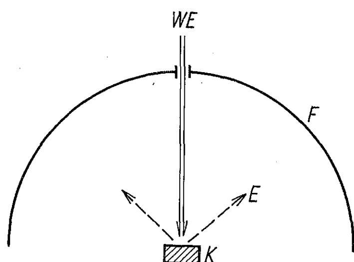
1/12
Obraz dyfrakcyjny kryształu krzemu otrzymany za pomocą wiązki elektronów. Zdjęcie wykonane na mikroskopie elektronowym w Instytucie Technologii Elektronowej Politechniki Warszawskiej

Doświadczenia prowadzone z elektronami o różnych energiach, jak również z innymi cząstkami wykazały, że zmierzone długości fali są zawsze zgodne z wartościami wynikającymi ze wzoru de Broglie'a.

Dla cząstek o większych masach zjawisko dyfrakcji nie może być obserwowane, ponieważ nie dysponujemy dostatecznie subtelnymi siatkami dyfrakcyjnymi, za pomocą których można by uzyskać obraz dyfrakcyjny.

W celu lepszego zrozumienia dualistycznej natury elektronów (i innych cząstek materialnych), przeanalizujmy dokładniej doświadczenie Davissona i Germera. W doświadczeniu tym bierze udział olbrzymia ilość elektronów. Nasuwa się więc pytanie, jak przebiegałoby ono, gdyby je wykonać na pojedynczym elektro-

nie. Przeprowadźmy w tym celu następujący eksperyment myślowy: Wyobraźmy sobie przedstawiony na rys. 1/13 przyrząd do badania dyfrakcji elektronów. Elektrony odbite od powierzchni kryształu padają na półkulistą kliszę fotograficzną, wywołując jej zaczernienie, podobnie jak to występuje w zwykłej fotografii, w której na kliszę jest rzutowany obraz optyczny. Lokalne zaczernienie kliszy w naszym doświadczeniu będzie proporcjonalne do liczby elektronów padających w określonym przedziale czasu na dany element kliszy.



1/13
Przyrząd do badania dyfrakcji elektronów
WE — wiązka elektronów padających na kryształ *K*; *E* — elektrony odbite; *F* — klisza fotograficzna

Załóżmy teraz, że zamiast wiązki elektronów stosowanej przez Davissona i Germera wprowądzamy do urządzenia pojedynczy elektron. Gdyby zachowywał się on tylko tak jak fala, to na kliszy otrzymalibyśmy taki sam obraz dyfrakcyjny, jak w doświadczeniu Davissona i Germera, z tą tylko różnicą, że jego intensywność (tzn. stopień zaczernienia kliszy) byłaby odpowiednio słabsza. W rzeczywistości tak się nie dzieje. Elektron ulega mianowicie odbiciu od powierzchni kryształu i pada na kliszę, powodując jej zaczernienie tylko w jednym punkcie. Powtarzając eksperyment na przykład 10-krotnie, otrzymalibyśmy 10 zaczernionych punktów rozsianych w sposób chaotyczny w różnych miejscach kliszy. Za każdym razem cała energia elektronu zostaje wydzielona w jednym punkcie kliszy, a więc mówiąc obrazowo, „cały" elektron zostaje tam „schwytany"

Zarówno fizyka klasyczna, jak i fizyka kwantowa nie są w stanie określić miejsc na kliszy, do których dolecą poszczególne elektrony. Fizyka kwantowa ma jednak nad fizyką klasyczną tę przewagę, że może dostarczyć bardzo ważnej informacji. Jak bowiem zobaczymy w następnym rozdziale, prawa fizyki kwantowej umożliwiają wyznaczenie *prawdopodobieństwa* dotarcia elektronu do dowolnie obranego obszaru kliszy.

Mamy tu więc do czynienia z sytuacją podobną do tej, jaką spotykamy na przykład przy rzucie kostką do gry. Przy pojedynczym rzucie nie jesteśmy w stanie określić jego wyniku. Jednocześnie wiemy jednak, że dla idealnie symetrycznej kostki prawdopodobieństwo wyrzucenia np. „piątki" wynosi dokładnie 1/6. Oznacza to, że jeżeli dokonamy dużej liczby (np. 6 milionów) rzutów jedną kostką, bądź też rzucimy jednorazowo 6 milionów kostek, to z bardzo dobrym przybliżeniem otrzymamy obrany wynik w 1 milionie przypadków[1].

---

[1] Oczekiwana niedokładność względna wynosi w tym przypadku 0,1% i dąży do zera w miarę powiększania liczby rzutów.

Podobnie przedstawia się sprawa z elektronami w doświadczeniu Davissona i Germera. Dla pojedynczego elektronu nie możemy przewidzieć miejsca zaczernienia kliszy. Jeżeli jednak dokonamy tego doświadczenia, wprowadzając do urządzenia, kolejno lub jednocześnie, bardzo wielką liczbę elektronów, to punkty zaczernione utworzą na kliszy obraz dyfrakcyjny pokazany na rys. 1/12. Rozkład zaczernienia kliszy można obliczyć z dużą dokładnością, traktując elektrony jako fale de Broglie'a, które ulegają odbiciu od siatki dyfrakcyjnej utworzonej przez kryształ.

Niewielka niedokładność obliczeń wiąże się ze skończoną liczbą użytych elektronów. W wyniku tego otrzymany obraz nie jest ciągły, lecz ma strukturę ziarnistą. Składa się on z wielkiej liczby dyskretnych punktów, z których każdy powstaje w wyniku wydzielenia się całej energii kinetycznej poszczególnych elektronów w emulsji fotograficznej[1].

Z opisanego powyżej doświadczenia myślowego można wysnuć następujące wnioski:

1. Elektrony mają jednocześnie własności korpuskularne i falowe. Własności te są indywidualnymi cechami każdego pojedynczego elektronu, a nie tylko, jak można by przypuszczać, dużego zbioru elektronów.

2. Rozchodzenie się elektronów w przestrzeni jest dobrze opisane przez równania falowe, dotyczące zachowania się fal de Broglie'a wraz z odpowiednimi warunkami brzegowymi, charakteryzującymi daną sytuację fizyczną. W szczególności elektrony mogą ulegać dyfrakcji i mogą interferować ze sobą, podobnie jak fale elektromagnetyczne.

3. Na podstawie rozkładu fal de Broglie'a można wyznaczyć prawdopodobieństwo wykrycia elektronu w danym obszarze, jeżeli umieści się tam odpowiedni detektor, np. w postaci kliszy fotograficznej.

4. Jeżeli w pewnym obszarze został „schwytany" elektron, to cała jego energia i cały ładunek zostają przekazane temu obszarowi.

Wnioski te odnoszą się nie tylko do elektronów (i innych cząstek materialnych), lecz także do fotonów. W tym przypadku równaniami opisującymi zachowanie się fotonów są równania Maxwella, a wielkością określającą gęstość prawdopodobieństwa wykrycia fotonu w danym punkcie jest gęstość energii pola elektromagnetycznego. Jest ona, jak to wynika ze wzoru (1/29), proporcjonalna do kwadratu amplitudy natężenia pola elektrycznego fali elektromagnetycznej. Na tej podstawie można oczekiwać, że kwadrat amplitudy fali de Broglie'a jest powiązany z prawdopodobieństwem znalezienia elektronu w danym obszarze.

Własności falowe elektronów są obecnie wykorzystywane m.in. w mikroskopach elektronowych. Urządzenia te, dzięki bardzo małej długości fali elektronów o dużej energii, charakteryzują się dużą zdolnością rozdzielczą. Umożliwiają one zatem uzyskiwanie obrazów o bardzo dużym powiększeniu.

Problem 1.4. Wyprowadź i przedyskutuj wzór uzależniający długość fali cząstki od jej energii kinetycznej wyrażonej w elektronowoltach (nie uwzględniaj efektów relatywistycznych).

---

[1] Abstrahujemy tu od ziarnistości samej emulsji, przyjmując, że jest ona nieskończenie drobnoziarnista.

Jaka jest długość fali elektronu w mikroskopie elektronowym o napięciu przyspieszającym 50 kV?

Problem 1.5. Aby zaobserwować dyfrakcję elektronów, należy użyć odpowiedniej siatki dyfrakcyjnej. Najsubtelniejszą siatką występującą w przyrodzie jest sieć krystaliczna. Davisson i Germer stosowali w swym doświadczeniu kryształ niklu. Używając elektronów o energii 54 eV, uzyskali oni odbicie braggowskie pierwszego rzędu (tzn. przy $n = 1$) pod kątem 46° w stosunku do wiązki padającej. Wyznacz odległość między płaszczyznami sieciowymi dla kryształu niklu.

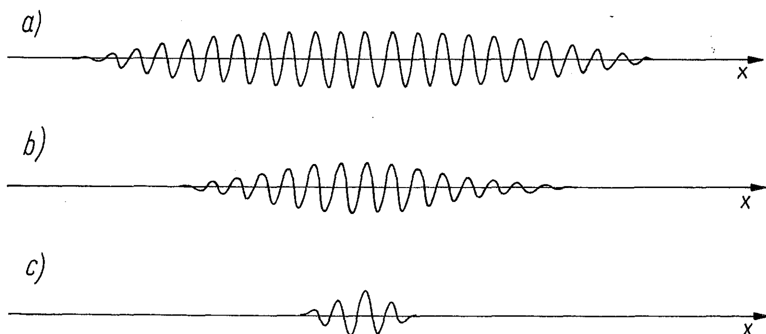
1.6 ZASADA NIEOZNACZONOŚCI

Opis ruchu cząstki w mechanice kwantowej różni się w zasadniczy sposób od opisu stosowanego w mechanice klasycznej. Zgodnie z założeniami mechaniki klasycznej, cząstka ma w każdej chwili określone położenie i określoną prędkość, a przez to również pęd. Są to wielkości, które zgodnie z prawami fizyki klasycznej, mogą być jednocześnie zmierzone z dowolnie dużą dokładnością.

W mechanice kwantowej nie istnieją ograniczenia dokładności, z jaką można wyznaczyć położenie cząstki, ani też dokładności, z jaką można zmierzyć jej pęd. Istnieje jednakże fundamentalne ograniczenie dokładności, z jaką można *jednocześnie* wyznaczyć położenie i pęd cząstki. Ograniczenie to jest konsekwencją falowej natury materii.

Pęd cząstki jest, jak to wynika ze wzoru (1/37), związany z długością przypisanej jej fali de Broglie'a. Cząstce o ściśle określonym pędzie odpowiadałaby zatem fala o ściśle określonej długości, tzn. fala monochromatyczna. Fala taka rozciągałaby się w przestrzeni w sposób nieograniczony, tzn. wypełniałaby całą przestrzeń. W wyniku tego położenie cząstki byłoby całkowicie nieokreślone.

Cząstce należy zatem przyporządkować paczkę falową, ograniczoną do pewnego obszaru przestrzeni. Im większa jest rozciągłość tej paczki, tym dokładniej jest określona długość fali, a więc również pęd cząstki, natomiast tym mniej dokładnie jest określone położenie cząstki. Kilka możliwych przypadków w tym zakresie przedstawiono w sposób schematyczny na rys. 1/14.



1/14 Paczki falowe reprezentujące cząstki kwantowe

Paczka falowa na rys. 1/14a obejmuje wiele cykli, a zatem długość fali i pęd cząstki mogą być wyznaczone z dużą dokładnością. Znajomość położenia cząstki jest natomiast niewielka, tzn. cząstka jest słabo zlokalizowana.

W przeciwieństwie do tego cząstka reprezentowana przez paczkę falową pokazaną na rys. 1/14c jest dobrze zlokalizowana. W tym przypadku trudno jest jednak mówić o długości fali, ponieważ rozpatrywany przebieg nie może być nawet w przybliżeniu uważany za okresowy.

Przedstawione tu idee znalazły swój wyraz w sformułowanej przez Heisenberga w 1927 r. *zasadzie nieoznaczoności*, w myśl której niedokładności Δx, Δy, Δz, z jakimi można wyznaczyć położenie cząstki, są związane z niedokładnościami Δp_x, Δp_y, Δp_z, z jakimi mogą być wyznaczone składowe jej pędu, następującymi zależnościami[1]:

$$\boxed{\begin{aligned} \Delta x\, \Delta p_x &\geqslant \hbar/2 \\ \Delta y\, \Delta p_y &\geqslant \hbar/2 \\ \Delta z\, \Delta p_z &\geqslant \hbar/2 \end{aligned}} \qquad (1/38)$$

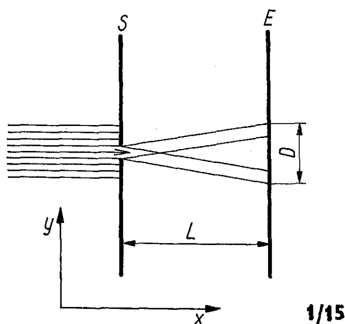
Z relacji Heisenberga wynika, że dokładność, z jaką może być znana współrzędna x jest zależna od dokładności ustalenia odpowiadającej jej składowej pędu p_x, nie zależy natomiast od dokładności określenia pozostałych składowych wektora pędu. Współrzędne położenia i pędu są więc skojarzone parami, przy czym dla każdej pary im dokładniej próbujemy określić jedną współrzędną, tym bardziej staje się nieokreślona druga współrzędna.

Nieoznaczoność ta nie wynika z niedoskonałości zastosowanych instrumentów pomiarowych, lecz z samej natury fizyki kwantowej. Skoro bowiem przyjmujemy falowy opis cząstek, to musimy dojść do wniosku, że odpowiednie współrzędne określające położenie i pęd cząstki nie mogą być jednocześnie znane z nieograniczoną dokładnością.

Problem 1.6. W celu zilustrowania konsekwencji zasady nieoznaczoności rozważmy zachowanie się elektronów przelatujących przez szczelinę (rys.1/1 5) o szerokości d. Załóżmy, że monoenergetyczna wiązka elektronów o energii W pada poprzez przesłonę S na ekran E umieszczony w odległości L. Położenie elektronu przelatującego przez szczelinę jest określone z dokładnością do $\Delta y = \pm d/2$. Wynika stąd pewna nieoznaczoność składowej p_y pędu elektronu, jak to uwidoczniono na rys. 1/15. Traktując elektrony korpuskularnie, powiedzielibyśmy, że tory elektronów ulegają odchyleniu, natomiast traktując je falowo, stwierdzamy że fala ulega ugięciu. W wyniku tego na ekranie pojawia się plamka o szerokości $D > d$. Wyprowadź i przedyskutuj wzór, umożliwiający oszacowanie szerokości plamki D w zależności od szerokości szczeliny d. Przy jakiej wartości d szerokość plamki staje się najmniejsza? Jaka jest w tych warunkach szerokość plamki? Wyznacz te wartości, przyjmując następujące dane: $W = 150$ eV, $L = 12{,}5$ cm.

---

[1] Najmniejsze wartości iloczynów $\Delta x\ \Delta p_x$ itd. równe $\hbar/2$ osiąga się dla paczek falowych, w których rozkłady współrzędnych położenia i pędu są rozkładami normalnymi (tzn. opisanymi funkcją Gaussa). Δx, Δp_x itd. oznaczają przy tym odchylenia standardowe odpowiednich współrzędnych.



1/15

W opisie kwantowomechanicznym cząstka (np. elektron lub foton) ma nie tylko pewną rozciągłość w przestrzeni (rys. 1/14), ale także w czasie. Odpowiednikiem długości fali charakteryzującej jej okresowość przestrzenną jest w dziedzinie czasu okres drgań. Jego odwrotność, tzn. częstotliwość jest związana z energią cząstki zależnością (1/36). Wynika stąd, że dla czasu i energii zachodzą analogiczne relacje nieoznaczoności, jak dla współrzędnych położenia i pędu. Znajduje to swój wyraz w następującej relacji Heisenberga:

$$\boxed{\Delta W \Delta\tau \geqslant \hbar/2} \tag{1/39}$$

gdzie ΔW oznacza niedokładność, z jaką jest określona energia cząstki, $\Delta\tau$ zaś — pewien czas charakterystyczny dla danego układu fizycznego (np. cząstki, atomu itd.).

Jeżeli układ znajduje się w stanie stacjonarnym, to w jego historii nie można wyróżnić żadnej chwili w stosunku do innych chwil. Należy więc przyjąć, że $\Delta\tau \to \infty$. Wynika stąd, że $\Delta W = 0$, co oznacza, że układ może mieć w tym stanie ściśle określoną energię.

Jeżeli natomiast układ znajduje się w stanie niestacjonarnym, to jako czas charakterystyczny należy przyjąć średni czas trwania tego stanu τ, tj. czas przejścia układu z jednego stanu stacjonarnego w inny stan stacjonarny. Nieokreśloność energii związanej z tym przejściem wyraża więc zależność

$$\Delta W \geqslant \frac{\hbar}{2\tau} \tag{1/40}$$

Problem 1.7. Wg obowiązującej obecnie definicji, metr jest to długość równa 1 650 563, 73 długości fali w próżni promieniowania odpowiadającego przejściu między poziomami $2p$ i $5d$ atomu kryptonu 86. Przejścia takie dają bardzo ostre linie widmowe, co oznacza, że towarzyszące im promieniowanie jest prawie monochromatyczne. Jeżeli jednak użyjemy przyrządu o dużej zdolności rozdzielczej, to można stwierdzić, że linie widmowe mają skończoną szerokość. Jedną z podstawowych przyczyn tego zjawiska jest to, że przejście i związana z nim emisja fotonu odbywają się w skończonym czasie. Zakładając, że czas ten jest rzędu 10^{-8} s, oszacuj względną szerokość linii odpowiadającej rozpatrywanemu przejściu. Co można sądzić o dokładności, z jaką została zdefiniowana jednostka długości?

2

ELEMENTY MECHANIKI KWANTOWEJ

Zjawiska opisane w poprzednim rozdziale wskazują na konieczność wprowadzenia do fizyki nowych koncepcji, których podstawą byłby obserwowany dualizm falowo-korpuskularny. Tym nowym podejściem do fal i cząstek materialnych jest *mechanika kwantowa*, której podstawy stworzyli w latach 1925–1926 m.in. Schrödinger, Heisenberg i Dirac. Teoria ta umożliwia wyjaśnienie struktury widm atomowych, istoty wiązań chemicznych oraz wielu innych zagadnień z zakresu fizyki atomowej i fizyki ciała stałego.

W rozdziale tym zaznajomimy się z podstawowymi postulatami mechaniki kwantowej w ujęciu Schrödingera i poznamy na kilku przykładach metody rozwiązywania najprostszych problemów dotyczących zachowania się elektronu w różnych sytuacjach fizycznych[1].

Mechanika kwantowa nie stanowi zaprzeczenia, ale pogłębienie i rozszerzenie mechaniki klasycznej, która daje poprawny opis zjawisk w przypadku cząstek o dużych rozmiarach i masach, zawodzi natomiast w przypadku mikrocząstek. Mechanika kwantowa jest więc teorią ogólniejszą, która przy przejściu od świata mikroskopowego do makroskopowego daje wyniki identyczne z fizyką klasyczną. Stwierdzenie to jest treścią tzw. *zasady korespondencji* (czyli zasady odpowiedniości) Bohra.

---

[1] Teoria Schrödingera jest najprostszą teorią mechaniki kwantowej. Jest to jednakże teoria nierelatywistyczna i w związku z tym nie daje ona możliwości ścisłego opisu wielu zjawisk zachodzących w świecie atomów. Dotyczy to np. oddziaływania promieniowania elektromagnetycznego z materią. Kontynuacją mechaniki kwantowej jest *elektrodynamika kwantowa,* która łączy teorię kwantów ze szczególną teorią względności.

2.1 FUNKCJA FALOWA

Odmienne własności elektronów i innych mikrocząsteczek w porównaniu z cząstkami „klasycznymi" wymagają nowego podejścia do opisu ich ruchu. W mechanice klasycznej pełny opis stanu cząstki punktowej uzyskuje się przez podanie w danej chwili jej 3 współrzędnych oraz 3 składowych wektora pędu. Znajomość tych danych początkowych umożliwia dokładne wyznaczenie przyszłego zachowania się cząstki w zadanym polu sił.

Opis tego rodzaju jest w mechanice kwantowej niemożliwy, ponieważ, zgodnie z zasadą nieoznaczoności, współrzędne cząstki i odpowiadające im składowe pędu nie mogą być jednocześnie znane z absolutną dokładnością. Opis kwantowy jest więc mniej precyzyjny niż opis klasyczny, a wskutek tego charakter przewidywań dotyczących zachowania się cząstki jest w mechanice kwantowej odmienny. Oznacza to, że jeżeli w mechanice kwantowej parametry ruchu elektronu zostały ustalone w możliwie najpełniejszy sposób, to i tak dalsze zachowanie się elektronu nie jest jednoznacznie określone. W mechanice kwantowej traci zatem sens pojęcie toru cząstki.

Zadanie mechaniki kwantowej sprowadza się do określenia *prawdopodobieństwa* otrzymania określonego wyniku pomiaru, np. prawdopodobieństwa tego, że elektron po upływie określonego czasu znajdzie się w określonym obszarze. Wyniki uzyskiwane w mechanice kwantowej mają więc charakter *statystyczny*, jak to już zaznaczono przy omawianiu doświadczenia Davissona i Germera w p. 1.5.

Opis ruchu w mechanice kwantowej musi uwzględnić falowe własności mikrocząstek, a zatem musi on być oparty na prawach rozchodzenia się fal. Zamiast równania ruchu, wyrażającego II zasadę dynamiki Newtona, występuje więc w mechanice kwantowej równanie falowe, zwane *równaniem Schrödingera.* Jego rozwiązaniem jest pewna funkcja współrzędnych i czasu

$$\Psi(x, y, z, t)$$

zwana *funkcją falową.* Jest to w ogólnym przypadku funkcja zespolona, która nie ma bezpośredniej interpretacji fizycznej i nie jest wielkością mierzalną.

Pomimo braku sensu fizycznego funkcja falowa odgrywa w mechanice kwantowej bardzo ważną rolę, określa ona bowiem całkowicie stan rozpatrywanej cząstki, tzn. zawiera w sobie wszystkie informacje o tym stanie. Znajomość tej funkcji umożliwia obliczenie prawdopodobieństwa otrzymania określonego wyniku dowolnego pomiaru, np. pomiaru położenia, pędu lub energii danej cząstki w rozpatrywanej chwili.

Bezpośredni sens fizyczny ma natomiast kwadrat modułu funkcji falowej. Określa on mianowicie prawdopodobieństwo znalezienia rozpatrywanej cząstki w elemencie przestrzeni o objętości $d\tau$ wokół punktu o współrzędnych x, y, z w chwili t. Prawdopodobieństwo to wynosi

$$|\Psi(x, y, z, t)|^2 \, d\tau$$

Stosunek tego prawdopodobieństwa do objętości $d\tau$ nazywamy *gęstością prawdopodobieństwa*. A zatem: *Kwadrat modułu funkcji falowej* $\Psi(x, y, z, t)$ *jest równy gęstości prawdopodobieństwa znalezienia danej cząstki w chwili* t *w punkcie o współrzędnych* x, y, z.

Z definicji prawdopodobieństwa wynika, że całka z $|\Psi|^2$ rozciągnięta na całą przestrzeń jest równa jedności

$$\iiint |\Psi|^2 \, d\tau = 1 \tag{2/1}$$

co oznacza, że znalezienie cząstki w całej przestrzeni jest pewne. Równanie (2/1) przedstawia tzw. *warunek unormowania* funkcji falowej.

Unormowana funkcja falowa jest określona z dokładnością do mnożnika $\exp(j\,\alpha)$, gdzie α jest dowolną liczbą rzeczywistą. Niejednoznaczność ta nie ma jednakże istotnego znaczenia, ponieważ, jak to już zaznaczono, funkcja falowa nie ma bezpośredniego sensu fizycznego i nie uwidocznia się w żadnym pomiarze fizycznym. Wielkością mierzalną jest jedynie kwadrat modułu tej funkcji, a pomnożenie funkcji falowej przez czynnik $\exp(j\,\alpha)$ nie zmienia jej modułu, wpływając tylko na wartość jej argumentu, tzn. na jej fazę. Z tego względu mnożnik $\exp(j\,\alpha)$ nazywany jest *czynnikiem fazowym*.

Funkcja Ψ i jej pochodne muszą być określone i ciągłe w całej przestrzeni, ponieważ prawdopodobieństwo znalezienia cząstki nie może być wielkością niejednoznaczną, nieskończoną lub zmieniać się w sposób skokowy.

2.2 RÓWNANIE SCHRÖDINGERA

Podane założenia dotyczące istnienia funkcji Ψ, jej regularności i interpretacji fizycznej kwadratu jej modułu stanowią podstawowe *postulaty* mechaniki kwantowej. Dalszym postulatem jest równanie Schrödingera, opisujące przebieg zmian stanu cząstki w czasie i w przestrzeni[1].

Jeżeli rozpatrujemy ruch cząstki o masie m w zadanym polu sił, to równanie to ma we współrzędnych prostokątnych następującą postać:

$$\frac{\hbar^2}{2m}\left(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}\right) - U\Psi = -j\hbar\frac{\partial\Psi}{\partial t} \tag{2/2}$$

gdzie $U = U(x, y, z, t)$ oznacza energię potencjalną charakteryzującą dane pole sił.

W wielu ważnych zagadnieniach mamy do czynienia z sytuacją, w której pole sił nie zmienia się w czasie (tzn. energia potencjalna U jest niezależna od czasu), a cząstka ma określoną energię całkowitą W. Funkcja falowa może być

---

[1] Równanie Schrödingera jest postulatem, tzn. tezą, którą się przyjmuje bez dowodu. Równania tego nie można więc wyprowadzić, można jedynie podać argumenty, które w pewnym stopniu wyjaśniają jego postać. Takie heurystyczne rozważania są zawarte np. w książce: R. Eisberg i R. Resnick: *Fizyka kwantowa atomów, cząsteczek, ciał stałych, jąder i cząstek elementarnych*. PWN, Warszawa 1983, p. 5.2.

wówczas przedstawiona w postaci iloczynu

$$\Psi(x, y, z, t) = \psi(x, y, z) \exp\left(-j\frac{W}{\hbar}t\right) \tag{2/3}$$

przy czym funkcja ψ, zwana *amplitudową funkcją falową*, zależy tylko od współrzędnych x, y, z.

We wzorze (2/3) jest zawarty postulat de Broglie'a, zgodnie z którym cząstce o energii całkowitej W przypisuje się częstotliwość ν określoną wzorem (1/36).

Na podstawie wzoru (2/3) można stwierdzić, że

$$|\Psi|^2 = |\psi|^2 \tag{2/4}$$

a zatem gęstość prawdopodobieństwa znalezienia cząstki w danym punkcie jest niezależna od czasu. W tych warunkach mówimy, że rozkład gęstości prawdopodobieństwa jest *stacjonarny*, a stany kwantowe spełniające ten warunek nazywamy *stanami stacjonarnymi.*

Podstawiając wyrażenie (2/3) do równania (2/2), otrzymujemy

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{2m}{\hbar^2}(W-U)\psi = 0 \tag{2/5}$$

W równaniu tym nie występuje czas, dlatego też jest ono nazywane *równaniem Schrödingera niezależnym od czasu.*

W dalszym ciągu tego rozdziału ograniczymy się do zagadnień jednowymiarowych, tj. takich, w których zarówno energia potencjalna, jak i funkcja falowa są zależne tylko od jednej współrzędnej x

$$U = U(x) \quad \text{oraz} \quad \psi = \psi(x)$$

Równanie Schrödingera upraszcza się wówczas do postaci

$$\boxed{\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(W-U)\psi = 0} \tag{2/6}$$

2.3 ELEKTRON SWOBODNY

Rozpatrzmy zachowanie się elektronu nie poddanego działaniu żadnych sił[1]. Bez ograniczenia ogólności rozważań możemy przyjąć, że $U = 0$. Całkowita energia elektronu jest wówczas energią kinetyczną, a zatem

$$W = \frac{p^2}{2m} \tag{2/7}$$

gdzie p oznacza pęd elektronu. Podstawiając (2/7) do równania (2/6), otrzymujemy

$$\frac{d^2\psi}{dx^2} + \frac{p^2}{\hbar^2}\psi = 0 \tag{2/8}$$

[1] W tym i w następnych punktach tego rozdziału będziemy rozpatrywali zachowanie się elektronu w różnych sytuacjach fizycznych. Rozważania te mają jednak charakter ogólny i odnoszą się do dowolnej cząstki o masie m.

Jest to równanie różniczkowe liniowe jednorodne o stałych współczynnikach. W książce tej spotkamy się wielokrotnie z równaniami o takiej postaci, dlatego też poświęcimy tu nieco uwagi metodzie ich rozwiązywania. Z zależności (2/8) wynika, że druga pochodna poszukiwanej funkcji $\psi(x)$ jest proporcjonalna do samej funkcji. Własność taką ma funkcja $\exp(\alpha x)$. Podstawiając ją do rozpatrywanego równania różniczkowego, otrzymuje się równanie algebraiczne

$$\alpha^2 + \frac{p^2}{\hbar^2} = 0 \tag{2/9}$$

zwane *równaniem charakterystycznym* danego równania różniczkowego. Pierwiastkami równania charakterystycznego są liczby α_1 i α_2. Ze względu na to, że równanie (2/8) jest liniowe, jego rozwiązanie ogólne jest kombinacją liniową funkcji $\exp(\alpha_1 x)$ oraz $\exp(\alpha_2 x)$, a więc ma postać

$$\psi(x) = A\exp(\alpha_1 x) + B\exp(\alpha_2 x) \tag{2/10}$$

gdzie A i B są to tzw. stałe dowolne, które wyznacza się z warunków początkowych lub brzegowych.

W rozpatrywanym tu przypadku pierwiastki równania charakterystycznego mają następujące wartości:

$$\alpha_1 = \mathrm{j}\frac{p}{\hbar} \quad \text{oraz} \quad \alpha_2 = -\mathrm{j}\frac{p}{\hbar}$$

zatem rozwiązanie ogólne równania (2/8) ma postać

$$\psi(x) = A\exp(\mathrm{j}kx) + B\exp(-\mathrm{j}kx) \tag{2/11}$$

ponieważ, zgodnie z zależnością (1/37) $\frac{p}{\hbar} = k$.

Otrzymane rozwiązanie przedstawia 2 fale o liczbie falowej k, rozchodzące się w dwóch przeciwnych kierunkach. Odpowiadają one elektronom poruszającym się w kierunku zgodnym i w kierunku przeciwnym do kierunku osi x. Ograniczając się do jednej z tych fal (odpowiadającej elektronowi poruszającemu się zgodnie z kierunkiem osi x) i podstawiając odpowiednie wyrażenie na $\psi(x)$ do wzoru (2/3), otrzymamy

$$\Psi(x, t) = A\exp\mathrm{j}(kx - \omega t) \tag{2/12}$$

Widzimy tu, że zarówno część rzeczywista, jak i część urojona funkcji $\Psi(x, t)$ propagują się w postaci fal harmonicznych wzdłuż osi x. Parametry tych fal, tzn. ich pulsacja i wektor falowy mają przy tym wartości wynikające z podanych w p. 1.5 wzorów de Broglie'a.

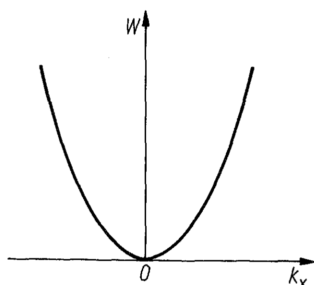
Problem 2.1. Wykaż, że funkcja (2/12) spełnia równanie Schrödingera o postaci (2.2).

W przypadku elektronu swobodnego rozwiązanie równania Schrödingera nie nakłada żadnych ograniczeń na wartość wektora falowego. Opierając się na wzorze (2/7), który można również przedstawić w postaci

$$W = \frac{\hbar^2}{2m}k^2 \tag{2/13}$$

należy stwierdzić, że nie istnieją również żadne ograniczenia kwantowomechaniczne na energię swobodnego elektronu. Podobnie jak w fizyce klasycznej, energia ta może więc przyjmować dowolną nieujemną wartość albo, inaczej mówiąc, elektron swobodny ma ciągłe widmo energetyczne.

Zależność wyrażoną wzorem (2/13) przedstawiono na rys. 2/1. Jak dalej zobaczymy, z takimi kwadratowymi zależnościami energii od wektora falowego będziemy mieli do czynienia również w bardziej skomplikowanych sytuacjach, np. w niektórych zakresach energii w półprzewodnikach.



2/1
Jednowymiarowy wykres przedstawiający zależność energii elektronu swobodnego od wektora falowego

Ze wzoru (2/12) wynika, że gęstość prawdopodobieństwa znalezienia elektronu

$$|\psi|^2 = A^2 \tag{2/14}$$

jest więc niezależna od współrzędnej x. Oznacza to, że jeżeli pęd elektronu jest dokładnie określony, to jego położenia nie można w ogóle wyznaczyć. Wniosek ten wynika zresztą bezpośrednio z zasady nieoznaczoności Heisenberga.

Chcąc unormować funkcję falową, należałoby narzucić warunek

$$\int |\psi|^2 \,\mathrm{d}x = 1 \tag{2/15}$$

który jednak nie może być spełniony dla żadnej skończonej, choćby dowolnie małej wartości A. Ta niemożność unormowania funkcji falowej wynika z założenia, że elektron może być reprezentowany przez idealnie monochromatyczną falę. W rzeczywistości, jak to już zaznaczono w p. 1.6, należy elektronowi przyporządkować paczkę falową, ograniczoną do pewnego obszaru przestrzeni. Wówczas unormowanie funkcji falowej staje się możliwe.

Paczka falowa reprezentująca elektron obejmuje pewien zakres częstotliwości wokół częstotliwości środkowej

$$\omega = \frac{W}{\hbar} = \frac{p^2}{2m\hbar} \tag{2/16}$$

Opierając się na tym wzorze, możemy wyznaczyć prędkość fazową i prędkość grupową paczki. Prędkość grupowa

$$v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{p}{m} = v \tag{2/17}$$

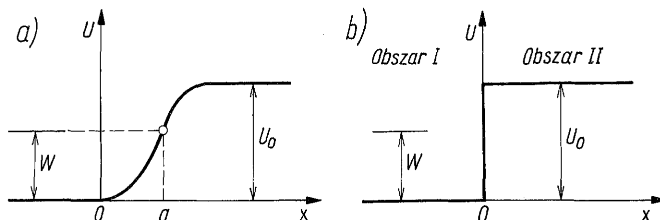
a więc paczka jako całość porusza się z prędkością równą prędkości elektronu v. Faza funkcji ψ przesuwa się natomiast wzdłuż osi x z prędkością

$$v_f = \frac{\omega}{k} = \frac{\hbar\omega}{\hbar k} = \frac{W}{p} = \frac{mc^2}{mv} = \frac{c^2}{v} \tag{2/18}$$

Jest to prędkość większa od prędkości światła, co jednak nie stoi w sprzeczności z teorią względności, ponieważ teoria ta stawia ograniczenie, jedynie dla prędkości, z jaką mogą być przenoszone sygnały i energia.

2.4 PRZEJŚCIE ELEKTRONU PRZEZ BARIERĘ POTENCJAŁU (PRZYPADEK $W < U_0$)

Rozpatrzmy zachowanie się elektronu w jednowymiarowym polu sił, w którym energia potencjalna U zmienia się wzdłuż osi x w sposób pokazany na rys. 2/2a. Elektron poruszający się w kierunku zgodnym z kierunkiem osi x natrafia więc na barierę potencjału o wysokości U_0. W punkcie tym zajmiemy się przypadkiem, gdy energia elektronu W jest mniejsza niż U_0.



2/2 Jednowymiarowa bariera potencjału: a) kształt rzeczywisty; b) kształt uproszczony

W zakresie $x < 0$ cała energia elektronu jest energią kinetyczną. Gdy elektron przekroczy punkt $x = 0$ zaczyna na niego działać siła hamująca, w wyniku czego jego energia potencjalna wzrasta, a energia kinetyczna maleje.

Przeanalizujmy najpierw zachowanie się elektronu wg praw mechaniki klasycznej. W miarę przesuwania się wzdłuż osi x elektron porusza się z coraz mniejszą prędkością i dociera do punktu $x = a$, w którym jego energia kinetyczna staje się równa zeru. W tym punkcie elektron zatrzymuje się i pod wpływem działającej na niego siły rozpoczyna ruch powrotny. Elektron zostaje więc odbity od bariery potencjału, ponieważ jego energia była niewystarczająca do pokonania tej bariery.

W celu przeanalizowania tego zagadnienia na podstawie równania Schrödingera przyjmijmy uproszczony kształt bariery potencjału przedstawiony na rys. 2/2b. Ta zmiana kształtu bariery nie ma wpływu na istotę zachodzących tu zjawisk, upraszcza jedynie obliczenia.

Rozwiązanie równania Schrödingera należy prowadzić oddzielnie dla obszaru I ($x \leqslant 0$) i dla obszaru II ($x \geqslant 0$), a następnie dopasować do siebie otrzymane rozwiązania tak, aby funkcja falowa i jej pochodna były ciągłe w punkcie $x = 0$. Operację tę nazywamy potocznie „*zszywaniem*" rozwiązań.

W obszarze I mamy $U(x) = 0$, a zatem równanie Schrödingera ma postać

$$\frac{d^2\psi_1}{dx^2} + \frac{2m}{\hbar^2} W\psi_1 = 0 \qquad (2/19)$$

przy czym ψ_1 oznacza amplitudową funkcję falową w tym obszarze. Równanie to rozwiązywaliśmy w p. 2.3, możemy więc od razu napisać jego rozwiązanie ogólne

$$\psi_1 = A \exp(jk_1 x) + B \exp(-jk_1 x) \qquad (2/20)$$

gdzie

$$k_1 = \sqrt{\frac{2mW}{\hbar^2}} \qquad (2/21)$$

Stałe dowolne A i B są to w ogólnym przypadku liczby zespolone. Jednakże ze względu na to, że funkcja falowa jest określona z dokładnością do czynnika fazowego, można przyjąć, że jedna z nich jest liczbą rzeczywistą. W dalszych rozważaniach założymy dla wygody, że stała A jest rzeczywista.

W wyrażeniu (2/20) pierwszy składnik przedstawia falę poruszającą się w kierunku zgodnym z kierunkiem osi x, tzn. falę *padającą*, drugi zaś – falę poruszającą się w przeciwnym kierunku, czyli falę *odbitą*.

W obszarze II $U(x) = U_0$, a zatem

$$\frac{d^2\psi_2}{dx^2} + \frac{2m}{\hbar^2}(W - U_0)\psi_2 = 0 \qquad (2/22)$$

Postępując analogicznie jak poprzednio, otrzymujemy

$$\alpha = \pm k_2 \qquad (2/23)$$

gdzie

$$k_2 = \sqrt{\frac{2m}{\hbar^2}(U_0 - W)} \qquad (2/24)$$

co prowadzi do następującego rozwiązania:

$$\psi_2 = C \exp(-k_2 x) + D \exp(k_2 x) \qquad (2/25)$$

Rozwiązania (2/20) i (2/25) różnią się zasadniczo. Jest to spowodowane tym, że w pierwszym obszarze równanie charakterystyczne ma pierwiaski urojone, w drugim zaś – rzeczywiste. W rezultacie tego funkcja ψ_1 jest kombinacją funkcji trygonometrycznych, a funkcja ψ_2 kombinacją funkcji wykładniczych.

W wyrażeniu (2/25) pierwszy składnik przedstawia funkcję, której moduł maleje asymptotycznie do zera, gdy $x \to \infty$, drugi zaś funkcję, której moduł rośnie nieograniczenie. Składnik ten należy odrzucić (tzn. przyjąć $D = 0$), ponieważ funkcja falowa nie może wzrastać w sposób nieograniczony. Uwzględniając to, otrzymujemy

$$\psi_2 = C \exp(-k_2 x) \qquad (2/26)$$

Przebieg ten traktuje się także jako falę, ale o urojonej wartości wektora falowego. Jest ona w optyce nazywana falą zanikającą lub falą *powierzchniową*,

gdyż występuje, praktycznie biorąc, jedynie w cienkiej warstwie przypowierzchniowej obszaru *II*.

Funkcja $\psi(x)$ i jej pochodna muszą być ciągłe w punkcie $x = 0$. Muszą być więc spełnione warunki dopasowania

$$\psi_1(0) = \psi_2(0) \tag{2/27}$$

$$\left.\frac{d\psi_1}{dx}\right|_{x=0} = \left.\frac{d\psi_2}{dx}\right|_{x=0} \tag{2/28}$$

Uwzględniając je, otrzymuje się następujące związki między stałymi dowolnymi A, B i C:

$$A+B = C \tag{2/29}$$

$$jk_1(A-B) = -k_2 C \tag{2/30}$$

Umożliwiają one wyznaczenie stałych B i C w zależności od stałej A, która pozostaje stałą dowolną – jest ona bowiem zależna od natężenia fali padającej. Z równań (2/29) i (2/30) wynika, że

$$B = \frac{jk_1+k_2}{jk_1-k_2} A \tag{2/31}$$

$$C = \frac{2jk_1}{jk_1-k_2} A \tag{2/32}$$

Korzystając z tych zależności oraz ze wzorów (2/20) i (2/26), otrzymuje się następujące wyrażenia na gęstość prawdopodobieństwa znalezienia elektronu w punkcie x:

dla $x \leqslant 0$ $(\psi = \psi_1)$

$$|\psi|^2 = 2A^2\left(1+\frac{k_1^2-k_2^2}{k_1^2+k_2^2}\cos 2k_1x - \frac{2k_1k_2}{k_1^2+k_2^2}\sin 2k_1x\right) \tag{2/33}$$

dla $x \geqslant 0$ $(\psi = \psi_2)$

$$|\psi|^2 = 4A^2\frac{k_1^2}{k_1^2+k_2^2}\exp(-2k_2x) \tag{2/34}$$

Podobnie jak w przypadku elektronu swobodnego, funkcja falowa wyrażona wzorami (2/33) i (2/34) nie może być unormowana. Elektron jest tu bowiem reprezentowany przez płaską falę harmoniczną, normalizacja zaś byłaby możliwa, gdyby potraktować go jako skończony ciąg falowy (paczkę falową).

Problem 2.2. Wykaż, że amplituda fali odbitej jest równa amplitudzie fali padającej, tzn. że $|A| = |B|$. Oznacza to, że elektron ulega całkowitemu odbiciu od bariery.

Wynik uzyskany w probl. 2/2 świadczy o tym, że w obszarze *I* opis kwantowomechaniczny zjawiska nie różni się w sposób istotny od opisu klasycznego. Zupełnie inaczej przedstawia się natomiast sprawa w obszarze *II*. Zgodnie z mechaniką klasyczną jest to dla elektronu obszar *zabroniony*, a zatem prawdopodobieństwo znalezienia elektronu w tym obszarze jest równe zeru. Analiza dokonana na podstawie równania Schrödingera prowadzi natomiast do wniosku,

że w tym klasycznie zabronionym obszarze występuje różna od zera funkcja falowa $\psi_2(x)$, a więc niezerowe prawdopodobieństwo znalezienia elektronu w dowolnie dużej odległości od progu bariery[1].

Jeżeli bariera potencjału ma skończoną grubość, to istnieje określone prawdopodobieństwo dotarcia elektronu do jej końca. Pojawia się więc fala przepuszczana przez barierę. Zjawisko to zwane *zjawiskiem tunelowym* będzie omawiane w p. 2.6[2].

Problem 2.3. Przedstaw za pomocą wykresu zależność $|\psi|^2$ od współrzędnej x dla elektronu o energii $W = \frac{1}{2} U_0$.

Rozwiązanie: Na podstawie wzorów (2/21) i (2/24) znajdujemy

$$k_1 = \sqrt{\frac{2m}{\hbar^2} W} = \sqrt{\frac{m}{\hbar^2} U_0}$$

$$k_2 = \sqrt{\frac{2m}{\hbar^2} (U_0 - W)} = \sqrt{\frac{m}{\hbar^2} U_0}$$

Wprowadzając oznaczenie $k_0 = \sqrt{\frac{m}{\hbar^2}} U_0$, można więc napisać

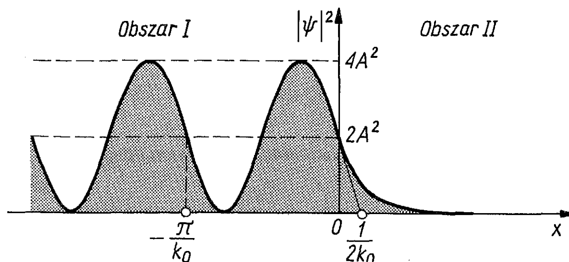
$$k_1 = k_2 = k_0$$

Podstawiając te wartości do wzorów (2/33) i (2/34) otrzymamy

$$|\psi_1|^2 = 2A^2(1 - \sin 2k_0 x)$$

$$|\psi_2|^2 = 2A^2 \exp(-2k_0 x)$$

Wykres tej zależności przedstawiono na rys. 2/3. W obszarze *I* występuje, jak widać, fala stojąca, powstająca w wyniku interferencji fali padającej i fali odbitej o jednakowych amplitudach. W obszarze *II* występuje fala o wartości malejącej wykładniczo.



2/3 Przebieg zależności $|\psi(x)|^2$ dla elektronu o energii $W = \frac{1}{2} U_0$ padającego na barierę o kształcie podanym na rys. 2/2b

Problem 2.4. Wyznacz odległość, na której gęstość prawdopodobieństwa znalezienia elektronu w obszarze *II* jest równa 1/100 wartości w punkcie $x=0$. Należy przyjąć dane jak w probl. 2.3 oraz $U_0 = 1$ eV. Jak odległość ta ma się do długości fali elektronu w obszarze *I*?

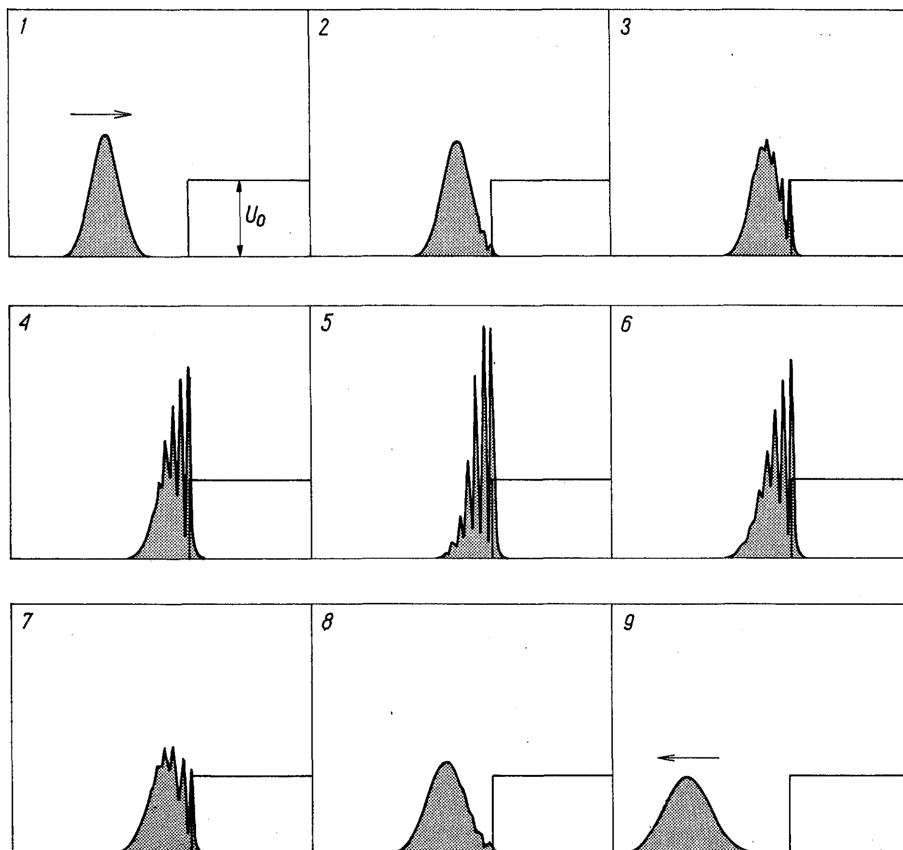
---

[1] Zjawisko odbicia elektronu od bariery potencjału i wnikania funkcji falowej do obszaru klasycznie zabronionego jest analogiczne do znanego z optyki zjawiska całkowitego wewnętrznego odbicia fali elektromagnetycznej, występującego w pewnych warunkach na granicy 2 ośrodków. W przypadku elektronu nie jest ono jednakże uwarunkowane kątem padania, lecz energią.

[2] Na wykorzystaniu tego zjawiska opiera się działanie niektórych rodzajów elementów półprzewodnikowych, np. diod tunelowych i diod Zenera.

Dokonana w tym punkcie analiza odbicia elektronu od bariery potencjału opiera się na rozwiązaniu równania Schrödingera niezależnego od czasu. Otrzymany wynik jest rozwiązaniem stacjonarnym, w którym giną główne elementy kinetyki tego zjawiska. Chcąc potraktować elektron jako paczkę falową, należy sięgnąć po równanie Schrödingera zależne od czasu. Podejście takie jest znacznie bardziej skomplikowane i wymaga zastosowania metod numerycznych oraz komputera o dużej mocy obliczeniowej.

Przykład rozwiązania tego zagadnienia dla jednowymiarowej paczki falowej przedstawiono na rys. 2/4. Zastosowano tu tzw. *paczkę gaussowską*, tzn. taką, w której rozkłady współrzędnych położenia i pędu są rozkładami normalnymi. Na kolejnych rysunkach przedstawiono przebieg energii potencjalnej charakteryzującej pole, w którym porusza się paczka oraz rozkład gęstości prawdopodobieństwa znalezienia elektronu. Jak można zaobserwować na rysunku, w miarę upływu czasu paczka zbliża się do bariery, wnika nieco w obszar klasycznie



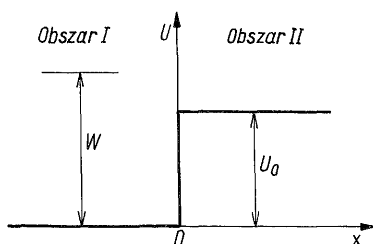
2/4 Odbicie gaussowskiej paczki falowej od prostokątnej bariery potencjału — średnia energia paczki równa połowie wysokości bariery
(Opracowano na podstawie: A. Goldberg, H. M. Schey: Computer-generated motion pictures of one-dimensional quantum-mechanical transmission and reflection phenomena. *Am. J. Phys.* **35**, 177 (1967))

zabroniony, a następnie zostaje całkowicie odbita od bariery. Skomplikowana struktura obwiedni paczki w czasie odbicia od bariery jest wywołana interferencjami fal padających i fal odbitych. Po odbiciu paczka przyjmuje na powrót kształt gaussowski, wykazując jednakże większą dyspersję niż na początku — odpowiednia krzywa przebiega, jak widać, bardziej płasko.

Wyniki przedstawione na rys. 2/4 oddają w sposób bardziej sugestywny rzeczywisty ruch elektronu. Podstawowe wnioski dotyczące zachowania się elektronu (całkowite odbicie od bariery oraz wnikanie w obszar klasycznie zabroniony) uzyskaliśmy jednakże już poprzednio, posługując się rozwiązaniami stacjonarnymi o postaci (2/20) i (2/26).

2.5* PRZEJŚCIE ELEKTRONU PRZEZ BARIERĘ POTENCJAŁU (PRZYPADEK $W>U_0$)

Rozważmy teraz przypadek, w którym energia elektronu W jest większa niż wysokość bariery potencjału U_0 — rys. 2/5. Według praw mechaniki klasycznej przejście elektronu z obszaru I do obszaru II jest absolutnie pewne. Jedynie energia kinetyczna elektronu ulega zmniejszeniu, wskutek tego, że przy przejściu przez płaszczyznę $x = 0$ część jej ulega przekształceniu w energię potencjalną. Elektron porusza się więc w obszarze II z mniejszą prędkością.



2/5
Bariera potencjału o kształcie prostokątnym

Rozwiązanie równania Schrödingera dla tego przypadku prowadzi do wniosku, że fala padająca ulega częściowemu odbiciu, częściowo zaś przechodzi do obszaru II i propaguje się w tym obszarze w postaci fali przepuszczonej. Zjawisko to występuje zawsze, niezależnie od tego, czy U_0 jest dodatnie, czy ujemne, a także niezależnie od różnicy $W-U_0$[1].

Problem 2.5. Przeanalizuj zachowanie się elektronu o energii W w polu, w którym energia potencjalna $U(x)$ zmienia się w sposób pokazany na rys. 2/5. Wyznacz funkcję falową $\psi(x)$ kwadrat jej modułu $|\psi(x)|^2$. Przedstaw za pomocą wykresu zależność $|\psi|^2$ od współrzędnej x dla elektronu o energii $W = 1,2\, U_0$.

Podobnie jak w poprzednich punktach, tak i w tym przypadku funkcja falowa nie może być unormowana. Elektron jest tu bowiem traktowany jako fala monochromatyczna. W rozpatrywanej sytuacji zwykle nie interesujemy się jed-

[1] Podobnie część energii fali elektromagnetycznej ulega zawsze odbiciu na granicy 2 ośrodków o różnych przenikalnościach elektrycznych.

nak gęstością prawdopodobieństwa znalezienia elektronu, lecz raczej natężeniem strumieni cząstek poruszających się w różnych obszarach i w różnych kierunkach. Ze wzorów (1/25) i (2/17) wynika, że każdy z tych strumieni jest proporcjonalny do iloczynu wektora falowego (w rozpatrywanym obszarze) i kwadratu amplitudy odpowiedniego składnika funkcji falowej.

Oznaczając podobnie jak w poprzednim punkcie przez $|A|$, $|B|$ i $|C|$ amplitudy fali padającej, odbitej i przepuszczonej, możemy więc zdefiniować współczynnik odbicia R jako

$$R = \frac{|B|^2}{|A|^2} \tag{2/35}$$

oraz współczynnik transmisji T jako

$$T = \frac{k_2|C|^2}{k_1|A|^2} \tag{2/36}$$

Między tymi współczynnikami zachodzi oczywisty związek

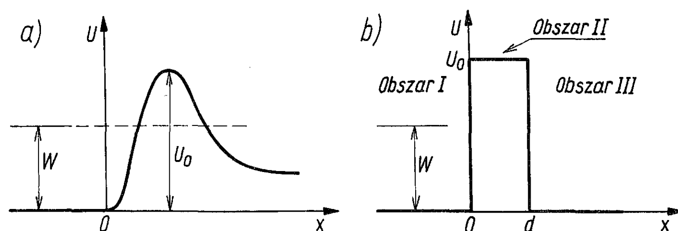
$$R+T = 1 \tag{2/37}$$

Problem 2.6. Sporządź wykres przedstawiający zależność współczynnika odbicia R od względnej wysokości bariery (tzn. od stosunku U_0/W) zarówno dla dodatnich, jak i dla ujemnych wartości U_0. Przedyskutuj otrzymane wyniki.

2.6* ZJAWISKO TUNELOWE

Zjawisko tunelowe polega na przenikaniu cząstek przez obszar klasycznie zabroniony, rozdzielający 2 obszary dozwolone (rys. 2/6). Zjawisko to jest efektem kwantowomechanicznym, nie mającym analogii w mechanice klasycznej. Zgodnie z prawami mechaniki klasycznej cząstka może pokonać barierę potencjału o wysokości U_0 tylko wówczas, gdy jej energia całkowita W jest większa niż U_0. W przypadku przedstawionym na rys. 2/6 następuje odbicie cząstki od bariery.

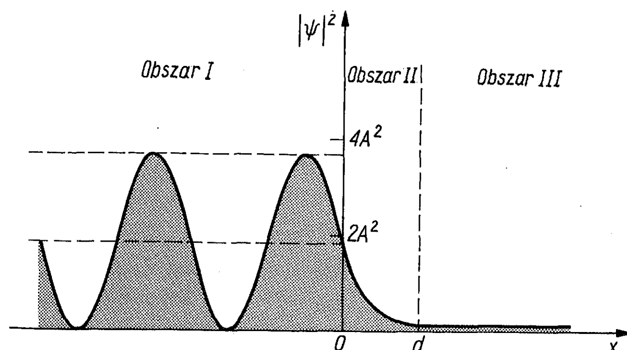
Dla cząstki kwantowej istnieje natomiast niezerowe prawdopodobieństwo przeniknięcia przez barierę. pomimo tego, że $W < U_0$. Przechodzi ona, jak gdyby tunelem na drugą stronę bariery i dlatego zjawisko to zostało nazwane *zjawiskiem tunelowym.*



2/6 Bariera potencjału o skończonej grubości: a) kształt rzeczywisty; b) kształt uproszczony

Zachowanie się elektronu padającego na barierę potencjału można przeanalizować, rozwiązując równanie Schrödingera. W celu uproszczenia rozważań, rozpatrzmy barierę o kształcie prostokątnym pokazanym na rys. 2/6b i, podobnie jak w probl. 2.5, przyjmijmy, że elektron ma energię $W = \frac{1}{2} U_0$.

Przebieg zależności $|\psi|^2$ od współrzędnej x dla tego przypadku przedstawiono na rys. 2/7. W obszarze *I* występuje, jak widać, fala stojąca, powstająca w wyniku interferencji fali padającej i fali odbitej od bariery. W rozważanym przypadku odbicie nie jest jednak całkowite, co oznacza, że amplituda fali odbitej jest mniejsza niż amplituda fali padającej.



2/7 Przebieg zależności $|\psi(x)|^2$ dla elektronu o energii $W = \frac{1}{2} U_0$ przechodzącego przez barierę o kształcie podanym na rys. 2/6b

W obszarze *II* występuje fala zanikająca – jest to bowiem obszar klasycznie zabroniony, a więc wektor falowy jest urojony.

W obszarze *III* występuje fala przepuszczona, której amplituda, podobnie jak dla elektronu swobodnego, jest niezależna od x.

Z rozwiązania równania Schrödingera wynika następujący wzór na współczynnik transmisji bariery prostokątnej:

$$T = \frac{1}{1+\eta^2 \operatorname{sh}^2(k_2 d)} \tag{2/38}$$

gdzie:

$$\eta = \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) \tag{2/39}$$

$$k_1 = \sqrt{\frac{2m}{\hbar^2} W} \tag{2/40}$$

$$k_2 = \sqrt{\frac{2m}{\hbar^2}(U_0 - W)} \tag{2/41}$$

Problem 2.7. Sporządź wykres przedstawiający zależność współczynnika transmisji T od wyrażenia $k_2 d$ dla elektronu o energii $W = \frac{1}{2} U_0$. Uprość wzór (2/38) dla $k_2 d \gg 1$. Przedyskutuj otrzymane wyniki.

Zjawisko tunelowe można wyjaśnić w sposób poglądowy za pomocą zasady nieoznaczoności Heisenberga. Cząstka klasyczna o energii $W_1 < U_0$ nie może znajdować się wewnątrz bariery potencjału, ponieważ jej energia kinetyczna W_k byłaby tam ujemna, a pęd urojony. Nie dotyczy to jednakże mikrocząstki, która zgodnie z zasadą nieoznaczoności nie ma jednocześnie ściśle określonego położenia i pędu. Ponieważ $W_k = p^2/2m$, zatem nieoznaczoność pędu jest związana z nieoznaczonością energii zależnością

$$\Delta W_k = \frac{2p\,\Delta p + (\Delta p)^2}{2m} \tag{2/42}$$

Przyjmując szacunkowo $\Delta p \approx p$, otrzymamy

$$\Delta W_k \approx \frac{3}{2m}(\Delta p)^2 \tag{2/43}$$

W celu wytłumaczenia przejścia cząstki przez barierę musimy przyjąć, że nieoznaczoność jej energii kinetycznej jest równa energii brakującej do pokonania bariery, tzn. że

$$\Delta W_k = U_0 - W_1 \tag{2/44}$$

Nieoznaczoność pędu wyniesie wówczas

$$\Delta p \approx \sqrt{\frac{2m}{3}(U_0 - W_1)} \tag{2/45}$$

a nieoznaczoność położenia cząstki

$$\Delta x \geqslant \frac{\hbar}{2\Delta p} \tag{2/46}$$

Jeżeli nieoznaczoność ta będzie równa grubości bariery d, to nie będziemy w stanie rozstrzygnąć, po której stronie bariery znajduje się cząstka. Warunek $\Delta x \approx d$ jest równoważny warunkowi

$$k_2\, d \approx \frac{\sqrt{3}}{2} \tag{2/47}$$

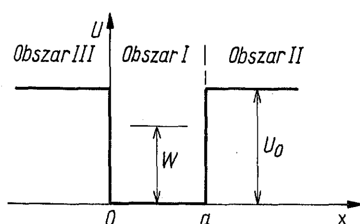
gdzie k_2 jest określone wzorem (2/41). Podstawiając tę wartość do wzoru (2/38) i przyjmując $\eta \approx 1$, otrzymujemy $T \approx 0{,}5$. Oznacza to, że w rozpatrywanym przypadku odbicie cząstki od bariery jest równie prawdopodobne, jak jej przeniknięcie przez barierę.

2.7 ELEKTRON W STUDNI POTENCJAŁU — STANY STACJONARNE

W poprzednich punktach tego rozdziału rozpatrywaliśmy ruch elektronu w warunkach, w których mógł on co najmniej w jednym kierunku oddalić się do nieskończoności. Stwierdziliśmy, że w tych warunkach energia elektronu nie jest skwantowana, tzn., podobnie jak w mechanice klasycznej, może mieć dowolną nieujemną wartość.

Obecnie rozważymy przypadek elektronu, którego ruch (przynajmniej w pewnym zakresie energii) jest ograniczony do obszaru skończonego. O takim elektronie mówimy, że znajduje się w *stanie związanym*. Jednym z najbardziej interesujących i ważnych przykładów takich sytuacji są stany elektronowe w atomie. Omówieniu ich jest poświęcony następny rozdział.

Jako wstęp do tych zagadnień rozpatrzymy teraz najprostszy możliwy przypadek stanu związanego. Zbadamy mianowicie zachowanie się elektronu w jednowymiarowym polu sił, w którym energia potencjalna zmienia się wzdłuż osi x w sposób przedstawiony na rys. 2/8. Założymy przy tym, że energia elektronu W jest mniejsza niż U_0. Mówimy wówczas, że elektron znajduje się w *studni potencjału.*



2/8
Elektron w jednowymiarowej studni potencjału

Zgodnie z prawami mechaniki klasycznej elektron jest uwięziony w obszarze ograniczonym współrzędnymi $x = 0$ i $x = a$ i niezależnie od posiadanej energii (przy założeniu, że $W < U_0$) wykonuje ruch wahadłowy, ulegając sprężystym odbiciom od ścianek studni. W celu znalezienia rozwiązania kwantowomechanicznego, należy rozwiązać równanie Schrödingera dla każdego z 3 obszarów zaznaczonych na rys. 2/8, a następnie dopasować do siebie uzyskane rozwiązania w taki sposób, aby funkcja falowa $\psi(x)$ i jej pochodna były ciągłe w punktach $x = 0$ oraz $x = a$.

W celu uproszczenia rozważań, założymy początkowo, że ścianki studni są nieskończenie wysokie, tzn. $U_0 \to \infty$. Wówczas dla $x < 0$ oraz $x > a$ gęstość prawdopodobieństwa znalezienia elektronu musi być równa zeru, a więc w obszarach *II* i *III* $\psi(x) = 0$. W obszarze środkowym równanie Schrödingera ma postać (2/19), a jego ogólnym rozwiązaniem jest funkcja

$$\psi = A \exp(\mathrm{j}\, kx) + B \exp(-\mathrm{j}\, kx) \tag{2/48}$$

gdzie

$$k = \sqrt{\frac{2m}{\hbar^2} W} \tag{2/49}$$

Z warunku ciągłości funkcji $\psi(x)$ w punkcie $x = 0$ wynika, że $A + B = 0$, a zatem w obszarze *I* funkcja falowa może być przedstawiona wzorem

$$\psi = A' \sin kx \tag{2/50}$$

Ciągłość funkcji falowej wymaga również, aby znikała ona w punkcie $x = a$. Nakłada to pewne ograniczenia na dopuszczalne wartości wektora falowego.

Wynikają one z warunku

$$A' \sin ka = 0 \tag{2/51}$$

co wymaga, aby

$$ka = n\pi \tag{2/52}$$

gdzie n jest liczbą naturalną.

Każdej wartości liczby n odpowiada więc określony stan kwantowy scharakteryzowany wektorem falowym

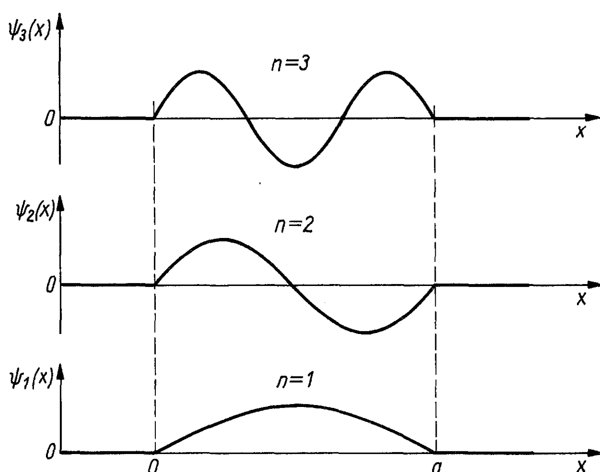
$$k_n = n\frac{\pi}{a} \tag{2/53}$$

i opisany amplitudową funkcją falową

$$\psi_n = A_n \sin\left(n\frac{\pi x}{a}\right) \tag{2/54}$$

gdzie A_n jest stałą, którą można wyznaczyć z warunku unormowania tej funkcji.

Liczbę n przyporządkowaną poszczególnym stanom nazywamy *liczbą kwantową*. Przebieg ψ_n dla 3 najmniejszych wartości n przedstawiono na rys. 2/9. Jak widzimy, liczba kwantowa n określa w rozpatrywanym przypadku liczbę połówek fali stojącej wypełniającej obszar, w którym uwięziony jest elektron.



2/9
Przebieg zależności ψ (x) dla trzech najniższych stanów energetycznych elektronu w prostokątnej studni potencjału o nieskończenie wysokich ściankach

Uważnego Czytelnika może zaniepokoić fakt, że w rozwiązaniach przedstawionych na rys. 2/9 pochodna funkcji falowej jest nieciągła w punktach $x = 0$ oraz $x = a$. Jest to spowodowane mało realnym założeniem nieskończenie wielkiego skoku energii potencjalnej w tych punktach. Przy dowolnie dużych, ale skończonych wartościach energii U_0 otrzymuje się przebiegi wykazujące ciągłość pochodnej funkcji falowej na brzegach obszaru I – zob. rys. 2/11.

Problem 2.8. Z warunku unormowania funkcji falowej wyznacz A_n, a następnie wykonaj wykres $|\psi_n(x)|^2$ dla 3 najmniejszych wartości liczby kwantowej n.

Każdemu stanowi kwantowemu odpowiada określona wartość energii elektronu, którą można wyznaczyć ze wzoru

$$W_n = \frac{\hbar^2 k^2}{2m} = n^2 \frac{h^2}{8ma^2} \tag{2/55}$$

Równanie Schrödingera dla elektronu znajdującego się w studni potencjału ma więc sensowne rozwiązanie stacjonarne o postaci

$$\Psi_n(x, t) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a} \exp\left(-\mathrm{j}\frac{W_n}{\hbar} t\right) \tag{2/56}$$

tylko wówczas, gdy energia elektronu przyjmuje jedną z wartości należących do dyskretnego zbioru W_1, W_2 ... W_n ..., określonego wzorem (2/55). Wszelkie inne wartości energii są wzbronione. Oznacza to, że dla innych wartości rozpatrywany problem nie ma fizycznie dopuszczalnego rozwiązania[1].

Skwantowanie energii elektronu uwięzionego w studni potencjału pojawiło się, jak widać, w sposób naturalny, jako konsekwencja falowych własności mikrocząstek. Uwidocznia się ono tym silniej, im mniejsze jest wyrażenie ma^2, a więc im mniejsza jest masa cząstki i im mniejszy obszar przestrzeni, w którym jest ona zlokalizowana. Przy wymiarach a znacznie przekraczających rozmiary atomowe różnice między dozwolonymi poziomami energetycznymi są tak nieznaczne, że w przybliżeniu można uważać, iż energia cząstki może zmieniać się w sposób ciągły. Jest to zgodne z cytowaną już zasadą korespondencji Bohra, która orzeka, że przy przejściu od skali atomowej do obiektów makroskopowych zależności kwantowe przechodzą w klasyczne.

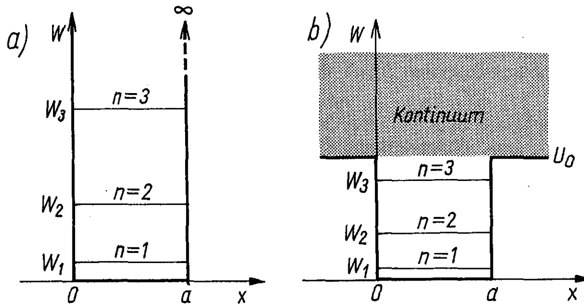
Jeżeli wysokość ścianek studni potencjału ma wartość skończoną, to gęstość prawdopodobieństwa znalezienia elektronu w obszarach, w których $U(x) = U_0$, jest różna od zera. Jak bowiem wynika z rozważań zawartych w p. 2.4, elektron może wnikać do obszarów klasycznie zabronionych. Podobnie jak w przypadku studni o nieskończenie wysokich ściankach, tak i w tym przypadku równanie Schrödingera ma sensowne rozwiązanie tylko dla pewnych wartości energii elektronu.

Problem 2.9. Dla elektronu znajdującego się w studni potencjału wyznacz ogólne rozwiązania równania Schrödingera w poszczególnych obszarach zaznaczonych na rys. 2/8. Znajdź równanie, z którego wynikają dopuszczalne wartości energii elektronu. Wykaż, że dla $U_0 \to \infty$ rozwiązanie tego równania ma postać przedstawioną wzorem (2/52).

Równanie otrzymane w probl. 2.9 jest równaniem przestępnym. Jego rozwiązanie nie może więc być przedstawione w postaci analitycznej. Przybliżone wartości pierwiastków równania można jednak dla każdego konkretnego przypadku wyznaczyć na przykład metodą numeryczną. Wyniki takich obliczeń przedstawiono na rys. 2/10b. Obok pokazano wyznaczone ze wzoru (2/55) wartości energii dla elektronu w studni o nieskończenie wysokich ściankach.

---

[1] Wartości energii W_n, dla których równanie Schrödingera ma sensowne rozwiązanie, nazywamy *wartościami własnymi* energii elektronu, a odpowiadające im funkcje falowe ψ_n — *funkcjami własnymi*.



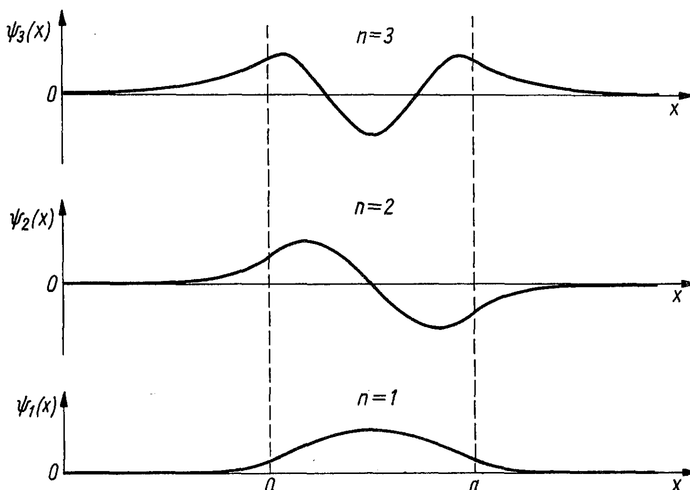
2/10 Widmo energetyczne elektronu w studni potencjału: a) studnia o nieskończenie wysokich ściankach; b) studnia o ściankach o wysokości U_0

Przy energii $W > U_0$ równanie Schrödingera ma rozwiązanie stacjonarne dla każdej wartości W. Energia elektronu nie jest więc w tych warunkach skwantowana, co wiąże się z faktem, iż elektron nie znajduje się w stanie związanym.

Dla studni o skończonej wysokości ścianek schemat poziomów energetycznych elektronu składa się więc ze zbioru dyskretnego o wartościach zawartych między zerem i U_0 oraz kontinuum rozciągającego się powyżej U_0. Początek kontinuum odpowiada energii, przy której układ rozpada się, co w naszym prostym przypadku oznacza, że z dala od obszaru centralnego elektron zachowuje się jak cząstka swobodna.

Przebiegi funkcji falowych dla stanów związanych z rys. 2/10b pokazano na rys. 2/11. W obszarze centralnym są to funkcje sinusowe, a poza tym obszarem funkcje wykładnicze zanikające.

Rozszerzymy teraz powyższe rozważania na zagadnienia trójwymiarowe. Rozpatrzymy przy tym najprostszy przypadek elektronu zamkniętego w tzw. *pudle potencjału*, tzn. w obszarze, w którym współrzędne x, y, z spełniają następujące warunki: $0 \leqslant x \leqslant a_x$, $0 \leqslant y \leqslant a_y$, $0 \leqslant z \leqslant a_z$.



2/11 Przebieg unormowanej funkcji ψ (x) dla stanów związanych elektronu w studni potencjału (rys. 2/10b)

Przyjmując, że w obszarze tym energia potencjalna elektronu jest równa zeru, a poza nim — nieskończenie wielka, można łatwo znaleźć rozwiązania równania Schrödingera. Można tu bowiem zastosować metodę rozdzielenia zmiennych, tzn. założyć, że poszukiwana funkcja falowa $\psi(y, x, z)$ może być przedstawiona w postaci iloczynu 3 funkcji, z których każda zależy tylko od jednej zmiennej

$$\psi(x, y, z) = \psi_1(x)\,\psi_2(y)\,\psi_3(z) \tag{2/57}$$

Równanie Schrödingera rozpada się wówczas na 3 niezależne równania różniczkowe zwyczajne, z których każde jest identyczne z równaniem dla jednowymiarowej studni potencjału. Trójwymiarowe warunki brzegowe sprawiają, że zamiast jednego równania (2/52) określającego dopuszczalne wartości wektora falowego, otrzymujemy 3 niezależne równania dla każdej z jego składowych. Stan kwantowy jest więc teraz wyznaczony przez 3 liczby kwantowe n_x, n_y i n_z, a dozwolone wartości energii elektronu są wyrażone wzorem

$$W = \frac{h^2}{8m}\left[\left(\frac{n_x}{a_x}\right)^2 + \left(\frac{n_y}{a_y}\right)^2 + \left(\frac{n_z}{a_z}\right)^2\right] \tag{2/58}$$

Ze wzoru tego wynika, że tej samej wartości energii mogą odpowiadać różne kombinacje liczb kwantowych, a więc również różne rozkłady przestrzenne gęstości prawdopodobieństwa znalezienia elektronu. O takich poziomach energetycznych, którym odpowiada kilka różnych stanów stacjonarnych mówi się, że są *zdegenerowane*. Możliwość istnienia tych stanów wynika z faktu, że w ogólnym przypadku sama energia nie określa w pełni stanu fizycznego cząstki.

Ze względu na to, że w trójwymiarowym pudle potencjału stan elektronu jest wyznaczony przez 3 liczby kwantowe, można oczekiwać, że również stany elektronowe w atomie będą w ujęciu Schrödingera opisane za pomocą trójki liczb kwantowych. Jak dalej zobaczymy, wniosek ten jest prawdziwy.

Problem 2.10. Dla elektronu zamkniętego w pudle potencjału o wymiarach $a_x = 2a_y = 3a_z$ wyznacz wartości liczb kwantowych dla 8 najniższych stanów energetycznych. Określ, które z nich są zdegenerowane.

2.8* ELEKTRON W STUDNI POTENCJAŁU — STANY NIESTACJONARNE

Równanie Schrödingera jest równaniem różniczkowym liniowym, a zatem dowolna kombinacja liniowa jego rozwiązań o postaci (2/56) stanowi nowe fizycznie dopuszczalne rozwiązanie. Weźmy zatem pod uwagę funkcję falową $\Psi(x, t)$, będącą superpozycją funkcji $\Psi_n(x, t)$ oraz $\Psi_m(x, t)$ opisaną wzorem

$$\Psi(x, t) = A_n\,\Psi_n(x, t) + A_m\,\Psi_m(x, t) \tag{2/59}$$

przy czym A_n i A_m są to współczynniki, które muszą być dobrane w taki sposób, aby funkcja $\Psi(x, t)$ była unormowana.

W celu uzyskania rozkładu prawdopodobieństwa znalezienia elektronu należy wyznaczyć kwadrat modułu funkcji $\Psi(x, t)$. Po podstawieniu odpowiednich wyrażeń do wzoru (2/59) i przeprowadzeniu obliczeń otrzymamy

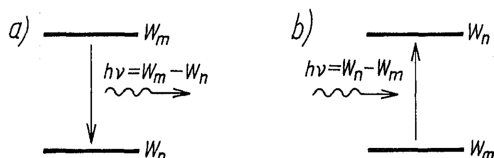
$$|\Psi(x, t)|^2 = A_n^2 \sin^2 \frac{n\pi x}{a} + A_m^2 \sin^2 \frac{m\pi x}{a} + \\ + 2A_n A_m \sin\frac{n\pi x}{a} \sin\frac{m\pi x}{a} \cos\left(\frac{W_m - W_n}{\hbar}\,t\right) \tag{2/60}$$

Jak łatwo możemy zauważyć, gęstość prawdopodobieństwa znalezienia elektronu w punkcie x jest teraz zależna od czasu. Mamy tu zatem do czynienia ze stanem *niestacjonarnym*, w którym energia elektronu zmienia się w czasie wskutek absorpcji lub emisji energii (np. w postaci promieniowania elektromagnetycznego). W wyniku tego elektron przechodzi z jednego stanu stacjonarnego w drugi, w rozpatrywanym przypadku np. ze stanu o energii W_m do stanu, w którym energia wynosi W_n.

Ostatni składnik we wzorze (2/60) wskazuje na to, że rozkład gęstości prawdopodobieństwa znalezienia elektronu wykonuje drgania harmoniczne o częstotliwości

$$\nu_{mn} = \frac{|W_m - W_n|}{\hbar} \tag{2/61}$$

Jeżeli $W_m > W_n$, to w wyniku tych drgań zostaje emitowany foton o energii równej różnicy $W_m - W_n$ i częstotliwości wyrażonej wzorem (2/61). Jeżeli natomiast $W_m < W_n$, to drgania te powstają wskutek absorpcji fotonu o energii równej $W_n - W_m$. Oba te przypadki pokazano na rys. 2/12.



2/12
Przejścia elektronu ze stanu o energii W_m do stanu o energii W_n: a) przejście z emisją fotonu; b) przejście z absorpcją fotonu

Różnice energii stanów stacjonarnych określają więc tzw. *częstotliwości rezonansowe* układu, tzn. częstotliwości promieniowania, które układ może emitować lub absorbować.

3

STANY ELEKTRONOWE W ATOMIE

Atom jest to najmniejsza cząstka pierwiastka zachowująca jego własności chemiczne. Składa się on z jądra, w którym skupiona jest prawie cała masa (ponad 99,9% masy atomu) i z poruszających się w jego otoczeniu elektronów.

Jądro atomowe jest zbudowane z protonów, obdarzonych dodatnim ładunkiem elektrycznym równym ładunkowi elementarnemu q, i z neutronów, pozbawionych ładunku elektrycznego. Liczbę protonów, wchodzących w skład jądra atomu danego pierwiastka, nazywamy *liczbą atomową* Z tego pierwiastka. Określa ona ładunek jądra (równy $+Zq$) oraz liczbę i konfigurację elektronów otaczających jądro. Liczba elektronów w elektrycznie obojętnym (tzn. nie zjonizowanym) atomie jest równa liczbie atomowej Z, a konfiguracja tych elektronów i związane z nią własności fizyczne i chemiczne danego pierwiastka zależą od liczby Z w sposób wykazujący prawidłowości, znajdujące swój wyraz w układzie okresowym pierwiastków Mendelejewa.

W rozdziale tym przedstawiono zarys współczesnej teorii budowy powłok elektronowych atomu. Teoria ta opisuje nie tylko zachowanie się elektronów w atomie, ale również umożliwia zrozumienie bardziej złożonych układów, takich jak cząsteczki chemiczne i ciała stałe.

Jako wstęp do rozważań, opartych na omawianych w poprzednim rozdziale prawach mechaniki kwantowej, przypomnimy poznaną w szkole średniej teorię atomu wodoru, określaną nazwą *modelu Bohra.*

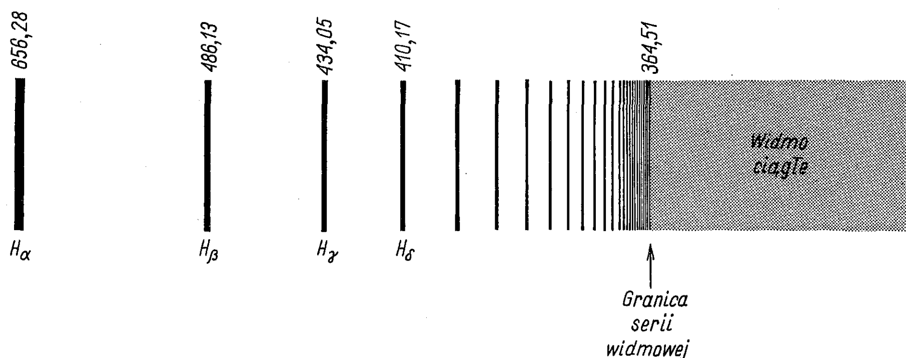
3.1 MODEL ATOMU BOHRA

Już w XIX wieku było znane dość dokładnie widmo emisyjne atomu wodoru, a w 1885 r. został opracowany przez Balmera wzór empiryczny dla linii tego widma. Wzór ten, uogólniony przez Rydberga, ma następującą postać:

$$\frac{1}{\lambda_{mn}} = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right) \qquad (3/1)$$

gdzie λ_{mn} – długość fali emitowanego promieniowania, m oraz n – liczby naturalne, przy czym $m > n$, R zaś – tzw. *stała Rydberga.*

Owczesna fizyka nie była w stanie wyjaśnić, dlaczego promieniowanie atomów nie ma widma ciągłego, lecz stanowi zbiór dyskretnych linii widmowych – rys. 3/1. Nie dawała ona również żadnych podstaw teoretycznych, umożliwiających otrzymanie wzoru Balmera.



3/1 Fragment widma emisyjnego atomu wodoru tzw. seria Balmera. Przy poszczególnych liniach podano ich oznaczenia oraz długość fali (w nanometrach)

W celu uzyskania zgodności teorii z doświadczeniem Bohr przyjął w 1913 r. pewne postulaty sprzeczne z fizyką klasyczną. Stworzony przez niego model atomu wodoru opiera się m.in. na założeniu, że elektron porusza się wokół jądra zgodnie z prawami mechaniki klasycznej, ale tylko po pewnych wyróżnionych orbitach kołowych, na których jego moment pędu jest całkowitą wielokrotnością stałej $\hbar$. Na tych dozwolonych orbitach musi więc być spełniony warunek

$$m_0 r v = n\hbar \qquad (3/2)$$

gdzie: r – promień orbity; v – prędkość elektronu; n – *liczba kwantowa* przyporządkowana danej orbicie.

Przy ruchu po okręgu koła zachodzi znany związek

$$m_0 \frac{v^2}{r} = F \qquad (3/3)$$

gdzie F oznacza siłę dośrodkową, która w rozpatrywanym przypadku jest siłą

elektrostatycznego przyciągania elektronu przez jądro. Zgodnie z prawem Coulomba

$$F = \frac{q^2}{4\pi \varepsilon_0 r^2} \tag{3/4}$$

Z tych 3 równań otrzymuje się następujący wzór na promienie dozwolonych orbit:

$$r_n = \frac{4\pi \varepsilon_0 \hbar^2}{m_0 q^2} n^2 \tag{3/5}$$

Rosną one więc w miarę zwiększania liczby n. Najmniejszą wartość promienia uzyskuje się dla $n = 1$. Wielkość tę nazywamy *promieniem Bohra* i oznaczamy symbolem a_0. Wynosi ona

$$a_0 = \frac{4\pi \varepsilon_0 \hbar^2}{m_0 q^2} = 5{,}29 \cdot 10^{-11}\ \text{m} = 52{,}9\ \text{pm} \tag{3/6}$$

Znając promienie orbit bohrowskich, można wyznaczyć dozwolone wartości energii elektronu w atomie wodoru. Energia ta jest sumą energii kinetycznej, związanej z ruchem elektronu po orbicie, i energii potencjalnej, wynikającej z oddziaływania jądra na elektron. Ta ostatnia jest, jak już mówiliśmy, określona z dokładnością do stałej addytywnej, tzn. jej wartość może być w pewnej odległości od jądra ustalona w sposób dowolny. Najdogodniej jest przyjąć, że energia potencjalna elektronu jest unormowana do zera w nieskończoności. Przy takim założeniu elektron, którego całkowita energia jest dodatnia, nie jest związany z jądrem (atom jest zjonizowany) i może poruszać się swobodnie w całej przestrzeni. Energia ta może mieć dowolną wartość, tzn. widmo energetyczne takiego elektronu jest ciągłe.

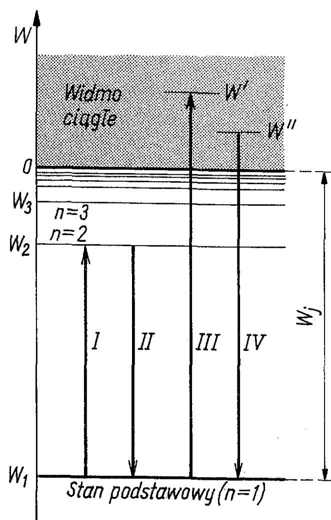
W przeciwieństwie do tego ujemnym wartościom całkowitej energii elektronu odpowiadają stany, w których jest on związany z jądrem i musi krążyć dokoła niego po jednej z dozwolonych orbit. Jego energia wyraża się wówczas wzorem

$$W_n = \frac{-q^2}{8\pi \varepsilon_0 a_0} \frac{1}{n^2} \tag{3/7}$$

Widmo energetyczne atomu wodoru przedstawiono na rys. 3/2. Podobnie jak dla elektronu poruszającego się w studni potencjału, składa się ono ze zbioru dyskretnych poziomów dla stanów związanych oraz z kontinuum odpowiadającego stanom swobodnym.

Dla $n = 1$ elektron znajduje się na orbicie położonej najbliżej jądra, a atom w najniższym stanie energetycznym zwanym *stanem podstawowym*. Atom nie podlegający działaniu sił zewnętrznych może w tym stanie przebywać przez czas nieograniczony.

Większym wartościom liczby n odpowiadają *stany wzbudzone*. Atom przechodzi ze stanu podstawowego do jednego ze stanów wzbudzonych (przejście *I* na rys. 3/2) wskutek absorpcji kwantu energii o wartości równej różnicy energii tych stanów. Energia ta może być dostarczona przez pochłonięcie fotonu (są to



3/2
Schemat poziomów energetycznych i przykłady przejść elektronowych w atomie wodoru: *I* — przejście ze stanu podstawowego na najniższy stan wzbudzony; *II* — przejście ze stanu wzbudzonego do stanu podstawowego; *III* — jonizacja atomu znajdującego się w stanie podstawowym; *IV* — rekombinacja elektronu o energii W'' ze zjonizowanym atomem

tzw. *przejścia optyczne*), bądź też w procesie zderzenia atomu z innym atomem lub elektronem.

Stany wzbudzone nie są trwałe, lecz charakteryzują się określonym, skończonym *czasem życia*, który dla większości z nich jest rzędu 10^{-8}–10^{-7} s. Ponieważ każdy układ fizyczny dąży do tego, by jego energia potencjalna miała najmniejszą możliwą wartość, zatem atom wzbudzony powraca samorzutnie do stanu podstawowego (w modelu Bohra elektron „przeskakuje" na orbitę znajdującą się najbliżej jądra), wypromieniowując foton o częstotliwości wyrażonej wzorem (2/61) – przejście *II* na rys. 3/2[1].

Jak widać na rys. 3/2, odstępy między kolejnymi poziomami energetycznymi zmniejszają się w sposób regularny. Odpowiednie poziomy energetyczne zagęszczają się więc i schodzą do granicy jonizacji, odpowiadającej oderwaniu elektronu od jądra. Odstęp między poziomem podstawowym a granicą jonizacji nazywamy *energią jonizacji* W_j danego pierwiastka.

Problem 3.1. Korzystając z modelu Bohra, wyprowadź wzór Balmera i wzór określający stałą Rydberga. Znajdź związek między tą stałą a energią jonizacji atomu wodoru. Wyznacz wartości liczbowe obu tych wielkości.

Przejście *III* na rys. 3/2 reprezentuje zjawisko jonizacji atomu znajdującego się w stanie podstawowym. Zjawisko to może zachodzić np. wskutek absorpcji fotonu o energii większej niż W_j – nosi ono wówczas nazwę *fotojonizacji*. Nad-

---

[1] Przejście do stanu podstawowego może odbyć się jednym „skokiem", bądź też (w przypadku $n > 2$) poprzez pośrednie stany wzbudzone. Atom emituje wówczas nie jeden, lecz odpowiednio większą liczbę fotonów. Atom może przekazać energię wzbudzenia również w inny sposób, tzn. bez emisji fotonu – są to tzw. przejścia *bezpromieniste*. Energia wzbudzenia jest w tym przypadku oddawana innej cząstce, np. innemu atomowi, z którym zderzył się rozpatrywany atom. Prawdopodobieństwo takich przejść jest zwykle znacznie mniejsze niż przejść promienistych.

wyżkę energii fotonu zachowuje elektron w postaci energii kinetycznej (energia W' na rys. 3/2), z którą oddala się on od jądra.

Zjawiskiem odwrotnym do jonizacji atomu jest rekombinacja elektronu ze zjonizowanym atomem (przejście *IV* na rys. 3/2). Zachodzi ona w wyniku zderzenia elektronu o energii kinetycznej W'' z jonem, wskutek czego atom przechodzi do stanu podstawowego, emitując foton o energii równej W_j+W''.

Widmo emisyjne atomu wodoru obliczone na podstawie modelu Bohra zgadza się dobrze z wynikami uzyskanymi doświadczalnie. Model ten może być jeszcze udoskonalony przez wprowadzenie orbit eliptycznych oraz przez uwzględnienie skończonej masy jądra i poprawki relatywistycznej, uzależniającej masę elektronu od jego prędkości.

Wszelkie próby natomiast uogólnienia modelu Bohra na atomy wieloelektronowe zakończyły się niepowodzeniem. Na podstawie modelu Bohra nie można również wyjaśnić, dlaczego atomy łączą się w cząsteczki. Zawodzi on nawet w tak prostym przypadku, jak jon H_2^+, w którym, podobnie jak w atomie wodoru, mamy do czynienia tylko z 1 elektronem.

Model Bohra odgrywa dziś już tylko historyczną rolę. Jest on sprzeczny z duchem mechaniki kwantowej, jest bowiem niezgodny z zasadą nieoznaczoności. Ponadto przypisuje on atomowi wodoru w stanie podstawowym różny od zera orbitalny moment pędu, podczas gdy wyniki pomiarów wykazują, że w stanie podstawowym moment ten jest równy zeru.

Przyczyną niepowodzenia teorii Bohra jest jej niespójność. Jest to teoria półklasyczna, w której do fizyki klasycznej zostały na siłę wprowadzone obce jej warunki kwantowe. We współczesnej teorii budowy atomów elektrony nie są traktowane jako cząstki, krążące po orbitach dokoła jądra, lecz jako paczki falowe poruszające się w otoczeniu jądra. W mechanice kwantowej nie jesteśmy bowiem w stanie określić toru elektronu, możemy tylko wyznaczyć prawdopodobieństwo przebywania elektronu w różnych obszarach.

3.2 KWANTOWOMECHANICZNA TEORIA ATOMU WODORU

Zapoznamy się teraz z teorią atomu wodoru opartą na rozwiązaniu równania Schrödingera. W celu uproszczenia rozważań przyjmiemy, że jądro atomowe jest nieskończenie małe, a jego masa nieskończenie wielka[1]. Jądro będziemy więc traktowali jako nieruchomy ładunek punktowy, wytwarzający pole elektryczne, w którym porusza się elektron. Jeżeli położenie jądra przyjmiemy za początek

[1] Promień jądra jest rzędu 10^{-13} cm, jest więc bardzo mały w porównaniu z rozmiarami atomu, którego promień jest rzędu $10^{-8}-10^{-7}$ cm. Masa elektronu jest 1840 razy mniejsza niż masa protonu lub neutronu. Można więc istotnie przyjmować, że w układzie współrzędnych związanych z atomem porusza się jedynie elektron, natomiast ruch jądra jest pomijalnie mały.

układu współrzędnych, to wytwarzane przez nie pole będzie opisane potencjałem kulombowskim

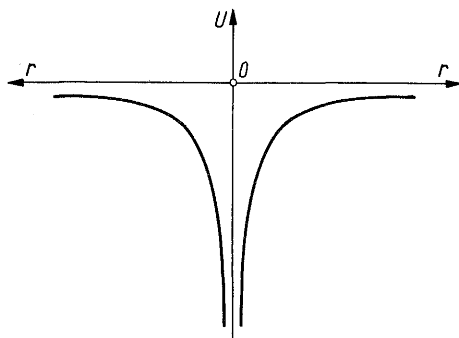
$$V(r) = \frac{q}{4\pi\,\varepsilon_0\, r} \tag{3/8}$$

Funkcja ta charakteryzuje się symetrią sferyczną. Oznacza to, że jej wartość w danym punkcie zależy tylko od odległości r tego punktu od początku współrzędnych, nie zależy natomiast od kierunku.

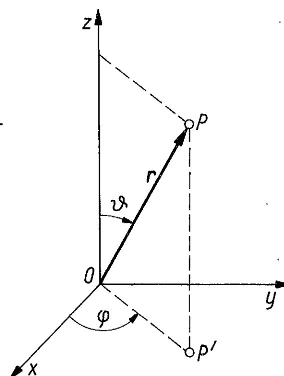
Znając potencjał, możemy napisać następujące wyrażenie na energię potencjalną elektronu w atomie wodoru:

$$U(r) = \frac{-q^2}{4\pi\,\varepsilon_0\, r} \tag{3/9}$$

Przebieg tej zależności pokazano na rys. 3/3. Ruch elektronu jest więc ruchem cząstki o 3 stopniach swobody w trójwymiarowej studni potencjału, której ścianki mają dla każdego kierunku kształt przedstawiony na rys. 3/3.



3/3 Zależność energii potencjalnej elektronu w atomie wodoru od jego odległości r od jądra



3/4 Układ współrzędnych sferycznych: $\mathbf{r}$ — wektor wodzący; ϑ — kąt biegunowy; φ — azymut

Ze względu na symetrię sferyczną pola sił, w którym porusza się elektron, dogodnie jest rozwiązywać równanie Schrödingera w układzie współrzędnych sferycznych pokazanym na rys. 3/4. Równanie to ma wówczas postać

$$\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{1}{r^2}\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\varphi^2}\right] +$$

$$+ \frac{2m_0}{\hbar^2}[W - U(r)]\,\psi = 0 \tag{3/10}$$

przy czym w rozpatrywanym przypadku jako $U(r)$ należy podstawić wyrażenie (3/9).

Problem 3.2. Znalezienie ogólnego rozwiązania równania (3/10) jest zadaniem trudnym. Z tego względu ograniczymy się tu do znalezienia jednego rozwiązania szczególnego. Nietrudno domyślić się, że ze względu na symetrię sferyczną pola sił, przynajmniej niektóre

rozwiązania będą miały także symetrię sferyczną, tzn. funkcja falowa ψ będzie zależała od promienia r, będzie natomiast niezależna od kątów ϑ i φ. Znajdź takie rozwiązanie.

Rozwiązanie: Przy założeniu symetrii sferycznej funkcji ψ pochodne $\frac{\partial\psi}{\partial\vartheta}$ i $\frac{\partial\psi}{\partial\varphi}$ znikają, a zatem równanie (3/10) upraszcza się do postaci

$$\frac{d^2\psi}{dr^2}+\frac{2}{r}\frac{d\psi}{dr}+\frac{2m_0}{\hbar^2}\left(W+\frac{q^2}{4\pi\varepsilon_0 r}\right)\psi=0 \qquad (3/11)$$

Jest to równanie liniowe jednorodne, ale występujące w nim współczynniki są zależne od zmiennej r. Przestawiając wyrazy tego równania, otrzymujemy

$$\left(\frac{d^2\psi}{dr^2}+\frac{2m_0}{\hbar^2}W\psi\right)+\frac{2}{r}\left(\frac{d\psi}{dr}+\frac{m_0 q^2}{4\pi\varepsilon_0\hbar^2}\psi\right)=0 \qquad (3/12)$$

Widoczne jest teraz, że warunkiem wystarczającym do tego, by równanie to było spełnione, jest znikanie obu wyrażeń w nawiasach. Prowadzi to do 2 równań o stałych współczynnikach, których sposób rozwiązania jest nam już dobrze znany. Jedno z możliwych rozwiązań równania (3/11) ma więc postać

$$\psi=A\exp(-kr) \qquad (3/13)$$

przy czym

$$k=\sqrt{-\frac{2m_0}{\hbar^2}W} \qquad (3/14)$$

Ponieważ dla elektronu w atomie $W<0$, wyrażenie pod pierwiastkiem jest dodatnie, a k – rzeczywiste[1].

Przyrównanie do zera drugiego składnika w równaniu (3/12) prowadzi do wniosku, że

$$k=\frac{m_0 q^2}{4\pi\varepsilon_0\hbar^2} \qquad (3/15)$$

lub inaczej

$$k=\frac{1}{a_0} \qquad (3/16)$$

przy czym a_0 jest to promień Bohra. Ma on teraz inną interpretację fizyczną niż w poprzednim punkcie, jest jednak zdefiniowany tym samym wzorem (3/6).

Eliminując k ze wzorów (3/14) i (3/15), otrzymujemy następujące wyrażenie na energię elektronu:

$$W=-\frac{m_0 q^4}{32\pi^2\varepsilon_0^2\hbar^2} \qquad (3/17)$$

Wzór ten jest równoważny ze wzoru (3/7) dla $n=1$. Sugeruje to, że znalezione rozwiązanie szczególne dotyczy stanu podstawowego atomu wodoru, i tak jest istotnie.

Problem 3.3. Korzystając z warunku unormowania funkcji falowej, wyznacz wartość współczynnika A we wzorze (3/13).

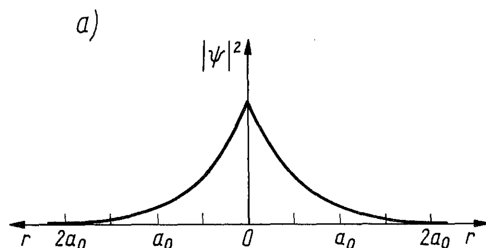
Wskazówka: Element objętości ma we współrzędnych sferycznych następującą postać:

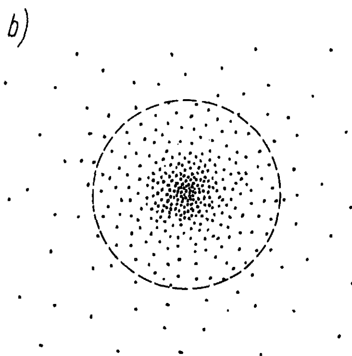
$$d\tau=r^2\sin\vartheta\,dr\,d\vartheta\,d\varphi \qquad (3/18)$$

Zależność kwadratu modułu funkcji falowej dla stanu podstawowego atomu wodoru od współrzędnej r przedstawiono na rys. 3/5a. Wielkość ta reprezentuje, jak wiadomo, gęstość prawdopodobieństwa znalezienia elektronu w danym punk-

---

[1] W wyrażeniu (3/13) założyliśmy z góry, że wykładnik jest ujemny, ponieważ dla dodatniego wykładnika funkcja falowa wzrastałaby w sposób nieograniczony.





3/5
Gęstość prawdopodobieństwa znalezienia elektronu w stanie podstawowym atomu wodoru: a) zależność $|\psi|^2$ od odległości r od jądra; b) model poglądowy chmury elektronowej (linia przerywana przedstawia przekrój powierzchni kuli o promieniu równym a_0)

cie. Jak wynika z rysunku, jest ona nierównomiernie rozłożona w przestrzeni i nie jest ograniczona do obszaru o skończonych wymiarach.

Z analizy rys. 3/5a wynika, że istnieje skończone prawdopodobieństwo znalezienia elektronu w obszarze bezpośredniego działania sił jądrowych. W wyniku tego powstaje możliwość wchłonięcia elektronu przez jądro atomowe. Zjawisko to, zachodzące w rzeczywistości i nazywane *wychwytem jądrowym elektronu*, jest procesem odwrotnym w stosunku do rozpadu promieniotwórczego beta. Procesowi wychwytu ulegają najczęściej elektrony z powłok wewnętrznych atomów pierwiastków ciężkich.

Ze względu na to, że chwilowe położenie elektronu w atomie nie jest ściśle określone, można elektron ten wyobrażać sobie w postaci rozmytej chmury przedstawionej w sposób poglądowy na rys. 3/5b. Gęstość punktów na tym rysunku obrazuje w przybliżeniu gęstość prawdopodobieństwa znalezienia elektronu w danym obszarze. Przyjmując taką interpretację funkcji falowej, często nazywamy wielkość $|\psi|^2$ *gęstością chmury elektronowej*, a wielkość $-q\,|\psi|^2$ — gęstością ładunku elektronowego.

Na podstawie rys. 3/5 nie należy jednak wyrobić sobie mylnego poglądu, że elektron w atomie ma postać nieruchomego rozmytego ładunku otaczającego jądro. Elektron porusza się bowiem i ma określoną energię i pęd, a także moment pędu, który w pewnych sytuacjach (np. w stanie podstawowym atomu wodoru) może być równy zeru. Rysunek 3/5 przedstawia po prostu rozkład uśrednionych w czasie położeń elektronu.

W wielu przypadkach dogodnie jest posługiwać się rozkładem gęstości prawdopodobieństwa uśrednionym nie tylko w czasie, lecz również po wszystkich kątach. Oznacza to, że interesuje nas prawdopodobieństwo znalezienia elektronu w danej odległości r od jądra. Prawdopodobieństwo to wyrażamy za pomocą

radialnej gęstości prawdopodobieństwa $P(r)$, zdefiniowanej w taki sposób, że prawdopodobieństwo znalezienia elektronu w warstwie sferycznej o promieniu r i grubości dr wynosi $P(r)\,dr$. Ponieważ objętość tej warstwy jest równa $4\pi r^2\,dr$, zatem dla funkcji falowej o symetrii sferycznej radialna gęstość prawdopodobieństwa znalezienia elektronu (czyli radialna gęstość chmury elektronowej) wyraża się wzorem

$$P(r) = 4\pi r^2|\psi(r)|^2 \tag{3/19}$$

Problem 3.4. Wykonaj wykres radialnej gęstości chmury elektronowej $P(r)$ dla stanu podstawowego atomu wodoru. Wyznacz najbardziej prawdopodobną odległość elektronu od jądra oraz średnią odległość elektronu do jądra. Przedyskutuj otrzymane wyniki.

Wskazówka: Wartość średnia (wartość oczekiwana, nadzieja matematyczna) zmiennej losowej nieciągłej przyjmującej wartości x_k $(k = 1, 2, ..., n)$ z prawdopodobieństwem p_k jest zdefiniowana wzorem

$$\langle x \rangle = \sum_{k=1}^{n} x_k p_k$$

Analogicznie dla zmiennej losowej ciągłej x o funkcji gęstości rozkładu $f(x)$ definiujemy wartość średnią następująco:

$$\langle x \rangle = \int x f(x)\,dx$$

przy czym całkowanie jest rozciągnięte na całą dziedzinę funkcji $f(x)$.

Powróćmy teraz do równania (3/10). Rozwiązuje się je metodą rozdzielenia zmiennych, tzn. zakłada, że poszukiwana funkcja falowa ψ ma następującą postać:

$$\psi(r, \vartheta, \varphi) = R(r)\,\Theta(\vartheta)\,\Phi(\varphi) \tag{3/20}$$

Podstawiając (3/20) do (3/10), otrzymuje się 3 niezależne równania różniczkowe zwyczajne, po jednym dla każdej ze współrzędnych r, ϑ, φ. Rozwiązania tych równań zawierają 3 parametry zwane *liczbami kwantowymi* i oznaczane symbolami n, l i m. Rozwiązania te mają sens fizyczny (tzn. funkcja ψ charakteryzuje się wymaganą regularnością i daje się unormować) tylko wówczas, gdy liczby kwantowe spełniają następujące warunki:

$$n = 1, 2, 3, ...$$
$$l = 0, 1, ... (n-1)$$
$$|m| = 0, 1, ... l$$

W teorii Schrödingera trójka liczb kwantowych n, l, m określa w sposób jednoznaczny stan ruchu elektronu w atomie. Liczby kwantowe wiążą się przy tym z najważniejszymi parametrami dynamicznymi elektronu poruszającego się w centralnym polu sił wytwarzanych przez jądro.

Liczba n, zwana *główną liczbą kwantową*, określa energię elektronu zgodnie ze wzorem (3/7) wyprowadzonym przez Bohra. Numeruje ona kolejne poziomy energetyczne stanów stacjonarnych atomu wodoru w porządku wzrastającej energii.

Każdej wartości liczby n odpowiada n możliwych wartości liczby l. Stany te mają tę samą energię, różnią się natomiast kształtem funkcji falowej i związa-

nym z nim orbitalnym momentem pędu elektronu. Z tego powodu liczba l nosi nazwę *orbitalnej liczby kwantowej*. Określa ona wartość momentu pędu elektronu $\boldsymbol{L}$ zgodnie ze wzorem

$$L = \sqrt{l(l+1)}\,\hbar \tag{3/21}$$

Z kolei każdej wartości liczby l odpowiada $(2l+1)$ możliwych wartości liczby m, zwanej *magnetyczną liczbą kwantową*. Określa ona składową momentu pędu elektronu dla wyróżnionego kierunku, zwanego kierunkiem *kwantyzacji*. Przyjmując ten kierunek za oś z, możemy napisać, że

$$L_z = m\hbar \tag{3/22}$$

Liczba ta ma istotne znaczenie tylko wówczas, gdy atom znajduje się w polu magnetycznym lub elektrycznym. Wówczas wyróżnionym kierunkiem jest kierunek tego pola. W przeciwnym przypadku, tzn. gdy atom nie jest poddany żadnym oddziaływaniom zewnętrznym, wszystkie kierunki są sobie równoważne, a zatem obierając którykolwiek z nich jako kierunek wyróżniony otrzymamy wartość składowej L_z zgodną ze wzorem (3/22). Z zasady nieoznaczoności wynika natomiast, że obie pozostałe składowe wektora $\boldsymbol{L}$ są wówczas całkowicie nieokreślone.

Dla zadanych wartości liczb kwantowych n, l, m funkcja falowa ma określoną postać, którą oznaczamy symbolem ψ_{nlm} i nazywamy *orbitalem atomowym*. Każdy orbital jest oznaczany tym samym symbolem co odpowiadający mu stan elektronowy.

Istnieją ogólnie przyjęte symbole, którymi oznacza się stany o różnych wartościach 2 pierwszych liczb kwantowych n i l. Tworzy się je z cyfry oznaczającej główną liczbę kwantową i litery przyporządkowanej liczbie l w następujący sposób:

$l = 0, 1, 2, 3, 4$

oznaczenie: $s\ p\ d\ f\ g$

Dla stanu podstawowego atomu wodoru $n = 1$, a zatem $l = m = 0$. Stan ten oznaczamy więc symbolem $1s$.

Dla najniższego stanu wzbudzonego $n = 2$, a zatem $l = 0$ lub 1. Dla $l = 0$ (stan $2s$) musi być również $m = 0$, natomiast dla $l = 1$ (stany $2p$) możemy mieć $m = -1$, 0 lub $+1$. W celu rozróżnienia stanów $2p$ o różnej wartości liczby m stosuje się niekiedy dodatkowy indeks podający wartość tej liczby.

Danej wartości energii mogą więc odpowiadać różne postacie funkcji falowej, czyli różne kształty orbitali atomowych. W rozpatrywanym przypadku $n = 2$ (tzn. $W = W_2$) istnieją 4 takie orbitale, które oznaczamy symbolami $2s$, $2p_{-1}$, $2p_0$ i $2p_{+1}$.

W tablicy 3/1 zestawiono stany kwantowe atomu wodoru dla $n = 1$ i 2, podając dla każdego z nich jego oznaczenie, wartości liczb kwantowych oraz oznaczenie funkcji falowej ψ_{nlm}. Przebiegi funkcji ψ_{nlm} dla $n = 1, 2, 3$ zestawiono w dodatku B.

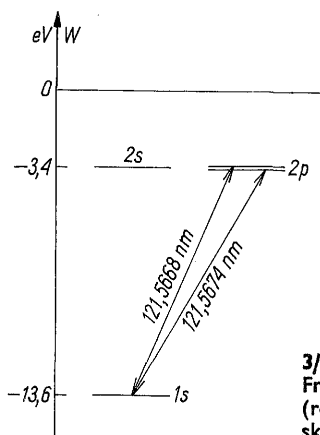
3/1
Stany kwantowe atomu wodoru dla 2 najmniejszych wartości liczby kwantowej n

Stan kwantowy	Liczby kwantowe			Funkcja falowa
	n	l	m	
$1s$	1	0	0	ψ_{100}
$2s$	2	0	0	ψ_{200}
$2p_0$	2	1	0	ψ_{210}
$2p_1$	2	1	1	ψ_{211}
$2p_{-1}$	2	1	-1	ψ_{21-1}

Problem 3.5. Uzupełnij tabl. 3/1 stanami kwantowymi dla $n = 3$. Wykaż, że dla danej wartości n głównej liczby kwantowej istnieje, zgodnie z teorią opartą na równaniu Schrödingera, n^2 różnych stanów kwantowych.

3.3 SPIN ELEKTRONU

Obserwacje widm atomowych potwierdzają z dość dobrym przybliżeniem kwantową teorię atomu opartą na równaniu Schrödingera. Na gruncie tej teorii nie można jednak wytłumaczyć tzw. *subtelnej struktury* linii widmowych, przejawiającej się w rozszczepieniu linii na kilka składników. Przykładowo, przejściom $2p \rightarrow 1s$ w atomie wodoru odpowiadają 2 linie widmowe o bardzo nieznacznie różniących się długościach fali – zob. rys. 3/6. Jest to spowodowane tym, że poziom $2p$ w tym atomie składa się w rzeczywistości z 2 bardzo blisko siebie położonych podpoziomów, różniących się o $5 \cdot 10^{-5}$ eV.



3/6
Fragment schematu poziomów energetycznych atomu wodoru (rozszczepienie poziomu $2p$ pokazano w bardzo powiększonej skali)

Bardziej doskonałą teorię atomu wodoru można opracować, rozwijając dalej teorię Schrödingera i przypisując elektronowi dodatkowy, czwarty stopień swobody. W teorii tej zakłada się mianowicie, że elektron ma pewien *własny* moment pędu, nie związany z jego ruchem w przestrzeni. Ten własny moment pędu elektronu nazywamy jego *spinem* w odróżnieniu od *orbitalnego* momentu pędu związanego z ruchem elektronu w polu sił jądra atomowego.

Spin jest specyficznie kwantową i zarazem relatywistyczną cechą cząstek elementarnych i w związku z tym nie może być wyjaśniony za pomocą modeli opartych na fizyce klasycznej. W szczególności nie jest poprawne traktowanie spinu jako momentu pędu wywołanego ruchem wirowym elektronu dokoła jego osi.

Przy uwzględnieniu spinu opis stanu elektronu powinien podawać nie tylko prawdopodobieństwo zajmowania przez niego określonego położenia w przestrzeni, lecz również prawdopodobieństwo orientacji jego spinu. Oznacza to, że funkcja falowa opisująca stan elektronu musi zależeć nie tylko od 3 współrzędnych przestrzennych, lecz również od czwartej zmiennej niezależnej zwanej *zmienną spinową*.

Do pełnego opisu stanu elektronu w atomie jest potrzebna zatem dodatkowa liczba kwantowa nazywana *liczbą spinową* i oznaczana symbolem m_s. Podaje ona wartość rzutu spinu na wybrany kierunek w przestrzeni (oś z)

$$S_z = m_s \hbar \qquad (3/23)$$

Doświadczenie pokazuje, że wiele cząstek elementarnych (np. elektrony, protony i neutrony) ma spin połówkowy, tzn. ich liczba spinowa może przyjmować jedną z 2 wartości: $m_s = 1/2$ lub $m_s = -1/2$. Cząstki takie podlegają statystyce Fermiego-Diraca i w związku z tym noszą nazwę *fermionów*[1]. Do cząstek tych odnosi się tzw. *zakaz Pauliego*. Jest to jedna z fundamentalnych zasad mechaniki kwantowej, głosząca, że w dowolnym układzie fizycznym 2 jednakowe cząstki o spinie połówkowym nie mogą jednocześnie znajdować się w tym samym stanie kwantowym, tzn. nie mogą mieć jednakowych wszystkich 4 liczb kwantowych.

Wprowadzenie spinu oraz zakazu Pauliego do mechaniki kwantowej umożliwiło dokładne wyjaśnienie struktury widm atomowych, istoty wiązań kowalencyjnych (występujących np. w cząsteczce H_2 oraz w kryształach diamentu i krzemu), jak również wielu innych zjawisk, których zrozumienie było poprzednio niemożliwe. Zakaz Pauliego stanowi także klucz do zrozumienia układu okresowego pierwiastków Mendelejewa.

3.4 ATOMY WIELOELEKTRONOWE

Teoria budowy atomów mających 2 lub więcej elektronów jest znacznie bardziej skomplikowana niż teoria atomu wodoru. Przyczyną tego jest konieczność uwzględ-

[1] Statystyka ta będzie omawiana w rozdz. 8

nienia oddziaływań występujących między elektronami. Największe znaczenie mają oddziaływania wynikające z elektrostatycznego odpychania się jednoimiennych ładunków. Znacznie słabsze są natomiast oddziaływania magnetyczne wywołane orbitalnym momentem pędu i spinem poszczególnych elektronów.

Równania uwzględniające wszystkie te zjawiska są bardzo skomplikowane, a ich analityczne rozwiązanie nie jest możliwe nawet w najprostszym przypadku atomu zawierającego 2 elektrony. Z tego powodu problem ten rozwiązuje się różnymi metodami przybliżonymi. W często stosowanej metodzie tzw. *przybliżenia jednoelektronowego* zaniedbuje się oddziaływania magnetyczne, a ponadto zakłada, że każdy elektron w atomie porusza się w uśrednionym polu elektrostatycznym wytworzonym przez jądro i wszystkie pozostałe elektrony. Przyjmuje się przy tym, że podobnie jak w atomie wodoru, pole to ma symetrię sferyczną. Jednakże, w odróżnieniu od atomu wodoru, nie jest ono opisane prostą zależnością wyrażoną wzorem (3/9), lecz jest bardziej złożoną funkcją promienia r.

W takim podejściu każdy elektron w atomie jest traktowany jako znajdujący się w swoim indywidualnym stanie kwantowym, określonym, podobnie jak w atomie wodoru, przez czwórkę liczb kwantowych: n, l, m i m_s.

W przybliżeniu jednoelektronowym energia elektronu zależy od liczb n i l, jest natomiast niezależna od m i m_s[1]. W stanie podstawowym atomu elektrony obsadzają najniższe możliwe stany energetyczne. Są to stany o najmniejszych wartościach liczb n i l.

Zbiór elektronów w atomie dzieli się na tzw. *powłoki*, zawierające elektrony o wspólnej głównej liczbie kwantowej n. Powłoki te oznacza się kolejno literami K, L, M, N, O, P i Q. Ze względu na zakaz Pauliego każda powłoka może przyjąć tylko ograniczoną liczbę elektronów. Dla powłoki utworzonej z elektronów o danej wartości liczby kwantowej n wynosi ona $2n^2$. Jest to wartość podwójna w stosunku do podanej w probl. 3.5, co wynika z tzw. *degeneracji spinowej*, tzn. 2 możliwych orientacji spinu dla każdego orbitalu określonego trójką liczb kwantowych n, l, m.

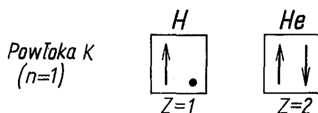
Rozpatrując zapełnianie się kolejnych powłok w atomach o coraz większej liczbie Z można wyjaśnić budowę układu okresowego pierwiastków[2].

Atom wodoru ($Z = 1$) składa się z jądra o ładunku $+q$ i 1 elektronu, który w stanie podstawowym atomu znajduje się w stanie $1s$. Atomem o liczbie atomowej $Z = 2$ jest hel. Jego jądro jest zbudowane z 2 protonów i 2 neutronów, a w polu sił jądra poruszają się 2 elektrony. Ze względu na 2 możliwe orientacje spinu oba elektrony mogą znajdować się w stanie $1s$. Mają one wówczas jednakowe wartości liczb kwantowych n, l i m, różnią się natomiast wartością liczby m_s. W atomie helu wszystkie stany o liczbie kwantowej $n = 1$ są obsadzone, a zatem pierwsza powłoka elektronowa, oznaczana symbolem K, jest całkowicie zapełniona. Pokazano to w sposób poglądowy na rys. 3/7.

---

[1] Dotyczy to warunków, w których atom nie jest poddany działaniu pola magnetycznego.

[2] Układ ten jest podany w dodatku C.

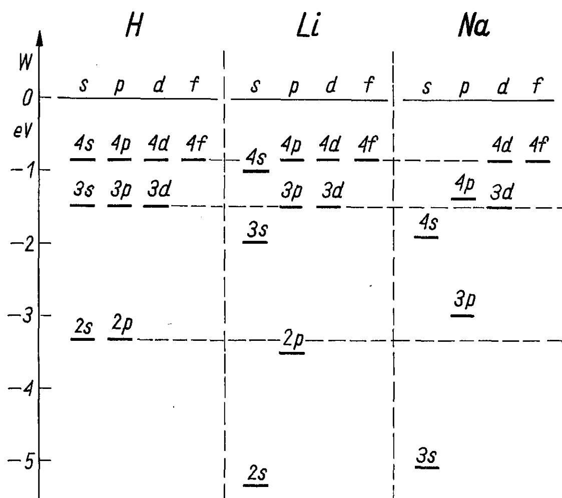


3/7 Zapełnianie się powłoki *K* w atomach tworzących I okres układu okresowego pierwiastków (kierunki strzałek odpowiadają orientacji spinu poszczególnych elektronów)

Zamknięta powłoka elektronowa, z jaką mamy do czynienia w atomie helu, charakteryzuje się dużą stabilnością. Atomy tego pierwiastka nie są więc skłonne ani do oddawania, ani też do przyjmowania elektronów. Z tego względu nie łączą się one ani ze sobą, ani też z atomami innych pierwiastków, są więc chemicznie nieczynne. Pierwiastki takie nazywamy *gazami szlachetnymi* i przypisujemy im wartościowość równą zeru.

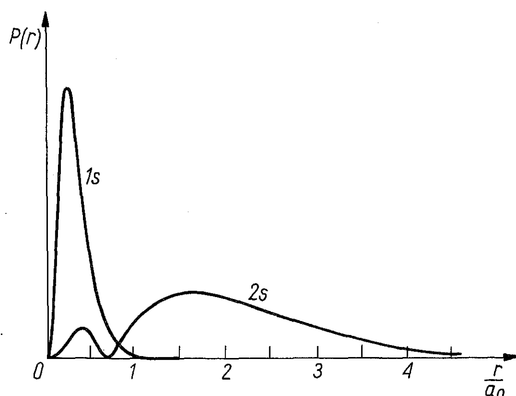
Wodór i hel tworzą pierwszy okres układu okresowego pierwiastków Mendelejewa. Drugi okres rozpoczyna się od litu ($Z = 3$). Spośród 3 elektronów atomu tego pierwiastka 2, podobnie jak w atomie helu, zapełniają powłokę *K*, trzeci natomiast, zgodnie z zakazem Pauliego, musi obsadzać wyższy poziom energetyczny o liczbie kwantowej $n = 2$. Rozpoczyna on więc zapełnianie powłoki *L*. Powłoka ta składa się z 2 podpowłok: podpowłoki $2s$ o 2 stanach kwantowych i podpowłoki $2p$ o 6 stanach kwantowych – zob. tabl. 3/1.

W atomie wodoru wszystkim stanom $2s$ i $2p$ odpowiada ta sama wartość energii ($W = W_2$), a więc poziom ten jest 8-krotnie zdegenerowany. W atomie litu natomiast, i ogólnie w atomach wieloelektronowych, energia danego stanu zależy, jak to już wyżej zaznaczono, nie tylko od głównej, lecz również od orbitalnej liczby kwantowej. Dla danej wartości liczby n rośnie ona ze wzrostem liczby l. Można to zaobserwować na rys. 3/8, na którym zestawiono poziomy energetyczne wodoru, litu i sodu – pierwiastka rozpoczynającego III okres układu okresowego.



3/8
Schemat poziomów energetycznych atomu wodoru, litu i sodu (dla atomu wodoru pominięto poziom 1s, dla pozostałych atomów pominięto poziomy całkowicie obsadzone)

Z rysunku tego wynika, że w stanie podstawowym atomu litu jego trzeci elektron znajduje się w stanie 2*s*. Rozkład radialnej gęstości prawdopodobieństwa znalezienia poszczególnych elektronów w tym atomie przedstawiono na rys. 3/9. Można na nim zauważyć, że elektrony, różniące się wartością głównej liczby kwantowej, tworzą istotnie odrębne powłoki, znajdujące się, w miarę powiększania *n*, w coraz większej odległości od jądra atomu. Z odległością tą wiąże się wartość energii potrzebnej do oderwania danego elektronu od atomu. Przykładowo, w atomie litu oderwanie elektronu 2*s* wymaga energii 5,4 eV, a do oderwania następnych elektronów konieczna jest energia wynosząca odpowiednio 75,6 i 122,4 eV. Elektrony 1*s* są więc w atomie litu bardzo silnie związane z jądrem i tworzą wraz z nim tzw. *rdzeń atomowy*, z którym dość luźno jest związany trzeci elektron.



3/9
Radialny rozkład prawdopodobieństwa znalezienia elektronu dla orbitali 1*s* oraz 2*s* atomu litu

Podobnie przedstawia się sprawa w atomach o większej liczbie elektronów. Elektrony tworzące wewnętrzne całkowicie zapełnione powłoki otaczają ciasno jądro i ekranują w znacznym stopniu jego dodatni ładunek. Na elektrony zewnętrznych niezapełnionych powłok działa zatem znacznie osłabiony ładunek, równy w przybliżeniu ładunkowi jądra zmniejszonemu o ładunek elektronów wchodzących w skład rdzenia atomowego.

Elektrony zewnętrznych podpowłok odłączają się więc stosunkowo łatwo od rdzenia atomowego. Jak zobaczymy w rozdz. 5, elektrony te biorą udział w tworzeniu wiązań w cząsteczkach chemicznych, jak również wiązań krystalicznych. Ich liczba i konfiguracja (tzn. rozkład na poszczególne podpowłoki) decydują o własnościach fizycznych i chemicznych pierwiastka, w tym przede wszystkim o jego *wartościowości*. Z tego względu noszą one nazwę *elektronów walencyjnych* lub elektronów wartościowości. Lit ma 1 elektron walencyjny, jest więc pierwiastkiem jednowartościowym.

W szeregu uporządkowanym według wzrastających liczb atomowych następnym pierwiastkiem jest beryl ($Z = 4$). Ma on 2 elektrony walencyjne, które w stanie podstawowym atomu obsadzają podpowłokę 2*s*. Ponieważ w podpowłoce tej mogą pomieścić się tylko 2 elektrony, zatem w kolejnych dalszych pierwiastkach następuje stopniowe zapełnianie podpowłoki 2*p*, która może pomieścić 6 elektronów.

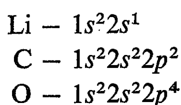
		Li	Be	B	C	N	O	F	Ne
		Z=3	Z=4	Z=5	Z=6	Z=7	Z=8	Z=9	Z=10
Powłoka L	2p			↑	↑↑	↑↑↑	↑↑↑↓	↑↑↑↓↓	↑↑↑↓↓↓
	2s	↑	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓
Powłoka K	1s	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓
Wartościowość:		+1	+2	+3	+4, −4	+5, −3	−2	−1	0

3/10 Rozkład elektronów na poszczególne powłoki i podpowłoki w atomach należących do II okresu układu okresowego pierwiastków. Przy każdym pierwiastku podano jego typową (ale nie jedyną) wartościowość. Liczby dodatnie oznaczają wartościowość względem tlenu (jest to pomnożona przez 2 liczba atomów tlenu, z jaką wiąże się atom danego pierwiastka). Liczby ujemne oznaczają wartościowość względem wodoru (jest to liczba atomów wodoru, z jaką wiąże się atom danego pierwiastka — dla rozpatrywanych pierwiastków jest to liczba elektronów potrzebnych do całkowitego zapełnienia powłoki *L*). Przykładowo azot ma wartościowość +5 (tworzy bowiem związek N_2O_5) lub −3 (ze względu na związek NH_3)

Konfiguracje elektronowe atomów należących do II okresu układu okresowego pierwiastków zestawiono na rys. 3/10[1]. Okres ten kończy się na neonie ($Z = 10$), w którym następuje całkowite zapełnienie powłoki *L*. Pierwiastek ten należy więc do grupy gazów szlachetnych.

Pierwiastkiem poprzedzającym neon jest fluor ($Z = 9$). W atomie fluoru brak jest 1 elektronu do całkowitego zapełnienia powłoki *L*. Z tego względu atomy tego pierwiastka wykazują duże *powinowactwo elektronowe*, tj. zdolność do przyłączenia z zewnątrz dodatkowego elektronu i tworzenia tym sposobem ujemnego jonu. Pierwiastkom takim przypisujemy wartościowość ujemną równą, co do wartości bezwzględnej, liczbie elektronów brakujących do zapełnienia zewnętrznej powłoki (lub podpowłoki) atomu. Fluor ma więc wartościowość równą −1.

Rozkład elektronów w atomie na poszczególne podpowłoki nosi nazwę *konfiguracji elektronowej* atomu. Rozkład ten podaje się, pisząc kolejne symbole podpowłok i zaznaczając cyframi u góry liczbę elektronów w każdej podpowłoce. Przykładowe oznaczenia konfiguracji niektórych atomów II okresu układu okresowego pierwiastków przedstawiają się w stanie podstawowym następująco:



Li – $1s^2 2s^1$
C – $1s^2 2s^2 2p^2$
O – $1s^2 2s^2 2p^4$

Często stosuje się oznaczenia uproszczone, w których pomija się wewnętrzne całkowicie zapełnione powłoki (lub podpowłoki), ograniczając się do podania konfiguracji elektronów zewnętrznych. Mówi się więc, że atom węgla ma konfigurację elektronową $2s^2 2p^2$.

---

[1] Jak wynika z rysunku, w przypadku 2 lub 3 elektronów w podpowłoce $2p$ spiny ich są ustawione równolegle. Jest to układ energetycznie korzystniejszy niż przy antyrównoległym ustawieniu spinów jednej pary elektronów.

W trzecim okresie, który rozpoczyna się od sodu ($Z = 11$), następuje kolejne zapełnianie podpowłok $3s$ i $3p$. Po zapełnieniu podpowłoki $3p$ rozpoczyna się nowy okres, ponieważ, jak to można zauważyć na schemacie poziomów energetycznych sodu (rys. 3/8), poziom $3d$ znajduje się powyżej poziomu $4s$. Zapełnianie powłoki $4s$, które rozpoczyna IV okres układu okresowego, następuje więc przy pustej podpowłoce $3d$.

W podpowłokach $3s$ i $3p$ można łącznie pomieścić 8 elektronów, a zatem III okres układu okresowego zawiera 8 pierwiastków i kończy się na argonie ($Z = 18$), który należy do grupy gazów szlachetnych.

Począwszy od IV okresu kolejność zapełniania się podpowłok elektronowych według wzrastających liczb kwantowych n, a w obrębie danej wartości liczby n – wg wzrastających liczb l, zostaje naruszona i zachodzi w przybliżeniu w następujący sposób:

IV okres $4s$, $3d$, $4p$
V okres $5s$, $4d$, $5p$
VI okres $6s$, $4f$, $5d$, $6p$
VII okres $7s$, $5f$, $6d$, $7p$

Jak można stwierdzić na podstawie tablicy podanej w dodatku C, od reguły tej istnieje kilka drobnych odstępstw, polegających na przesunięciu jednego lub obu elektronów z najbardziej zewnętrznej podpowłoki s do wewnętrznych podpowłok d lub f.

Podany tu schemat zapełniania się powłok ogranicza liczbę elektronów najbardziej zewnętrznej powłoki do 8. Liczba elektronów walencyjnych różnych pierwiastków może więc zawierać się w granicach od 0 do 7.

W miarę wzrostu liczby atomowej Z powtarzają się okresowo konfiguracje elektronów walencyjnych. Przykładowo, pierwsze 3 pierwiastki IV grupy układu okresowego mają następujące konfiguracje elektronowe (oznaczenia uproszczone):

C – $2s^2 2p^2$
Si – $3s^2 3p^2$
Ge – $4s^2 4p^2$

To podobieństwo konfiguracji zewnętrznych elektronów tłumaczy okresowe powtarzanie się własności fizycznych i chemicznych pierwiastków.

4

STRUKTURA KRYSTALICZNA CIAŁ STAŁYCH

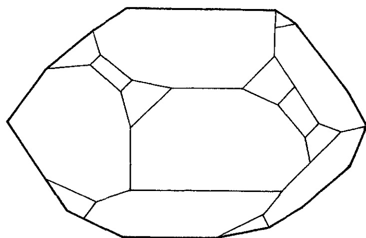
W poprzednim rozdziale zapoznaliśmy się z budową powłok elektronowych atomu. Atomy mogą, dzięki występującym między nimi siłom przyciągania, łączyć się w cząsteczki chemiczne (molekuły) albo w zespoły, zawierające olbrzymie ilości atomów, które, w zależności od stanu skupienia, zaliczamy do ciał stałych lub cieczy. Przedmiotem dalszych rozważań w tej książce będą własności elektronowe ciał stałych, dlatego też w rozdziale tym zaznajomimy się z budową tych ciał.

4.1 CIAŁA KRYSTALICZNE

Ciało stałe definiuje się ogólnie jako ciało odznaczające się stabilnością kształtu. W zależności od stopnia uporządkowania struktury wewnętrznej dzielimy ciała stałe na *krystaliczne* i *amorficzne*, czyli bezpostaciowe.

Ciała krystaliczne mają strukturę uporządkowaną. Oznacza to, że w ich budowie istnieje określona prawidłowość ułożenia atomów w obszarach bardzo dużych w porównaniu z rozmiarami atomu. Taką regularność struktury nazywamy *uporządkowaniem dalekiego zasięgu*. Jeżeli uporządkowanie to dotyczy całej bryły materiału, to stanowi ona pojedynczy kryształ, zwany *monokryształem* – rys. 4/1. Jeżeli natomiast bryła materiału składa się z chaotycznie ułożonych drobnych ziaren krystalicznych, to materiał ten określamy jako *polikrystaliczny*.

Proces krystalizacji ciała stałego rozpoczyna się zwykle jednocześnie w wielu miejscach, dlatego też większość spotykanych w przyrodzie ciał krystalicznych (np. minerały, metale itd.) ma budowę polikrystaliczną. Monokryształy służące do badań fizycznych i do zastosowań technicznych są zwykle wytwarzane sztucznie (jest to tzw. *hodowanie* kryształów). Proces ten wymaga szczególnej czystości materiałów wyjściowych oraz zachowania określonych warunków krystalizacji.



4/1
Naturalny monokryształ kwarcu (SiO_2)

Największe, wytwarzane na skalę przemysłową monokryształy krzemu mają średnicę 15,5 cm. Długość ich przekracza 2 m, a masa dochodzi do kilkudziesięciu kilogramów.

Jedną z cech charakterystycznych kryształu jest jego *anizotropia*, polegająca na tym, że ośrodek ten wykazuje w różnych kierunkach różne własności fizyczne.

W odróżnieniu od kryształów, ciecze i ciała amorficzne (których typowym przedstawicielem jest szkło) nie mają prawidłowej struktury krystalicznej i najczęściej nie wykazują anizotropii. W ciałach tych występuje jedynie tzw. uporządkowanie *bliskiego zasięgu*, polegające na tym, że w stosunku do dowolnego atomu rozkład najbliższych atomów jest w znacznym stopniu uporządkowany, chociaż nie tak dokładnie jak w kryształach. Jednak w miarę oddalania się od danego atomu prawidłowość ułożenia innych atomów w stosunku do niego zmniejsza się i przy odległościach rzędu 1 nm zwykle zanika zupełnie[1)].

Ciała krystaliczne charakteryzują się ponadto tym, że przejście z fazy stałej w fazę ciekłą następuje u nich w pewnej ściśle określonej temperaturze, nazywanej temperaturą topnienia. W przeciwieństwie do czego, w ciałach amorficznych przejście to zachodzi w sposób stopniowy, wskutek czego właściwie nie istnieje granica między tymi fazami. Z tego względu ciała amorficzne są często traktowane jako przechłodzone ciecze.

Ciało stałe, będące zespołem złożonym z dużej liczby atomów, może być opisane na gruncie wiedzy o strukturze pojedynczego atomu. Kwantowa teoria kryształów jest obecnie dobrze rozwinięta, a własności fizyczne tych materiałów zostały dobrze poznane. Teoria ciał amorficznych znajduje się natomiast dopiero w początkowym stadium rozwoju. W książce tej będziemy się zajmowali jedynie monokrystalicznymi ciałami stałymi, których własności zostały najlepiej poznane i które we współczesnej elektronice mają dominujące znaczenie.

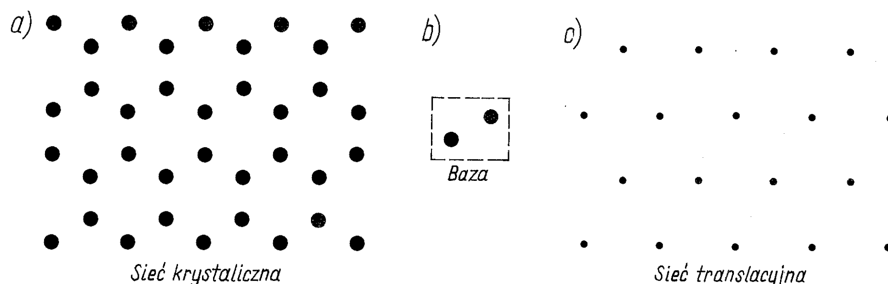
4.2 SIEĆ KRYSTALICZNA

Idealny nieskończony kryształ jest strukturą utworzoną przez regularne powtarzanie w przestrzeni identycznych elementów. Ten charakterystyczny dla danego ciała układ atomów nazywamy jego *siecią krystaliczną*.

[1)] Ciała, które zazwyczaj mają postać krystaliczną, np. metale i krzem, mogą być również wytwarzane w postaci amorficznej. Przykładowo, krzem amorficzny jest obecnie powszechnie stosowany w bateriach słonecznych.

Najmniejszy pojedynczy element sieci, powtarzający się periodycznie w 3 wymiarach przestrzeni, nazywamy *bazą* tej sieci. W najprostszym przypadku (np. w kryształach sodu, miedzi, złota i wielu innych metali) baza jest utworzona z 1 atomu danego pierwiastka. Na ogół jednak zawiera ona 2 lub więcej atomów, przy czym mogą to być atomy tego samego lub różnych pierwiastków. Przykładowo, w krysztale diamentu bazę stanowią 2 atomy węgla, w krysztale soli kuchennej 1 atom sodu i 1 atom chloru, ale na przykład w krysztale boru baza, w zależności od odmiany alotropowej, zawiera od 12 do 190 atomów.

Grupy atomów stanowiące bazę są ustawione w krysztale wg określonej sieci przestrzennej, którą nazywamy *siecią translacyjną* lub *siecią punktową*. Jest ona utworzona z punktów nazywanych *węzłami* sieci. Przykład sieci krystalicznej oraz jej bazę i sieć translacyjną przedstawiono na rys. 4/2. W celu ułatwienia zobrazowania, obrano tu sieć dwuwymiarową. Czytelnik może jednak z łatwością rozszerzyć rozważania zawarte w tym punkcie na sieci trójwymiarowe występujące w realnych kryształach.



4/2 Dwuwymiarowa sieć krystaliczna (a), jej baza (b) i sieć translacyjna (c)

Jeżeli przez dowolne 2 węzły sieci translacyjnej przeprowadzi się prostą, to znajdzie się na niej nieskończenie wiele węzłów położonych w jednakowych od siebie odległościach[1]. Prostą taką nazywamy *prostą sieciową*.

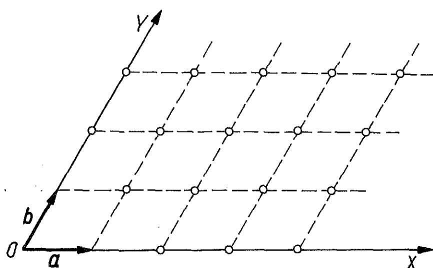
Nieograniczona sieć punktowa wykazuje tzw. *symetrię translacyjną*, tzn. przekształca się w siebie w wyniku przesunięcia równoległego (translacji) w kierunku dowolnej prostej sieciowej o dowolną wielokrotność odcinka nazywanego *okresem identyczności* lub *odcinkiem translacji*. Ujmując to krócej, można powiedzieć, że w sieci translacyjnej wszystkie kierunki i odległości propagują się w całej przestrzeni.

Symetria translacyjna sieci wyraża się tym, że położenie dowolnego węzła w stosunku do węzła przyjętego za punkt odniesienia może być (dla sieci dwuwymiarowej) określone za pomocą wektora

$$\boldsymbol{R} = m\boldsymbol{a} + n\boldsymbol{b} \qquad (4/1)$$

gdzie: m, n – liczby całkowite, $\boldsymbol{a}, \boldsymbol{b}$ zaś – odpowiednio dobrane wektory, nazy-

[1] Własności tej nie ma natomiast sieć przedstawiona na rys. 4/2a, dlatego też musi ona być traktowana jako sieć z dwuatomową bazą.



4/3
Dwuwymiarowa sieć translacyjna
X, Y — osie krystalograficzne, *a*, *b* — elementarne wektory translacji

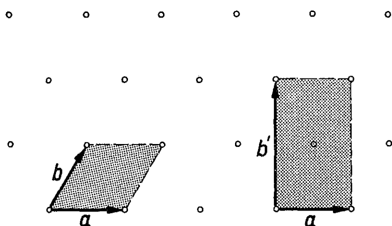
wane *elementarnymi wektorami translacji* (rys. 4/3). Długości tych wektorów, tzn. liczby *a* i *b* nazywamy *stałymi sieci*.

Osie *X*, *Y*, na których leżą wektory ***a***, ***b*** nazywamy *osiami krystalograficznymi*. Tworzą one układ współrzędnych, który w ogólnym przypadku jest układem ukośnokątnym. Wybór wektorów ***a***, ***b*** i związanych z nimi osi krystalograficznych może być dokonany w różny sposób. Zwykle jednak dąży się do tego, aby kąt zawarty między osiami różnił się jak najmniej od 90°.

Równoległobok zbudowany na elementarnych wektorach translacji nazywamy *komórką elementarną prostą*, lub prymitywną. Jak wynika z rys. 4/3, sieć przestrzenną kryształu stanowi nieskończony zbiór jednakowych komórek ułożonych równolegle względem siebie i wypełniających całą przestrzeń. W komórce prostej węzły sieci translacyjnej znajdują się tylko w narożach, a zatem na 1 komórkę przypada 1 węzeł.

W niektórych rodzajach sieci dogodnie jest posługiwać się komórkami elementarnymi o większych rozmiarach, ale za to o wyższym stopniu symetrii niż komórka prosta. Są to tzw. *komórki złożone*, które oprócz węzłów w narożach zawierają dodatkowe węzły (w przypadku sieci dwuwymiarowej – jeden dodatkowy węzeł w środku komórki). Sieci takie nazywamy sieciami *centrowanymi*[1].

Przykładem dwuwymiarowej sieci centrowanej jest sieć przedstawiona na rys. 4/3. Na rys. 4/4 pokazano jej komórkę prostą i komórkę złożoną. Komórka prosta jest równoległobokiem ukośnokątnym o 1 elemencie symetrii (środek symetrii), natomiast komórka złożona jest prostokątem, ma zatem 3 elementy symetrii (środek i 2 osie symetrii).



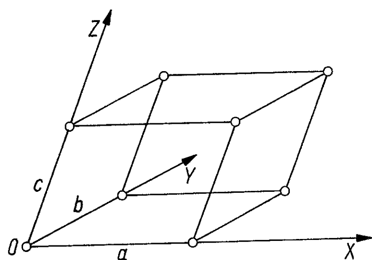
4/4
Komórki elementarne prosta i złożona sieci prostokątnej centrowanej

[1] W sieci trójwymiarowej dodatkowy węzeł możne znajdować się w środku komórki (sieć przestrzennie centrowana) albo też w środku wszystkich lub niektórych ścian (sieci płasko centrowane).

Problem 4.1. Istnieje 5 rodzajów dwuwymiarowych sieci translacyjnych, jednym z nich jest przedstawiona na rys. 4/4 sieć prostokątna centrowana. Narysuj pozostałe 4 rodzaje i nadaj im stosowne nazwy. Wyjaśnij, dlaczego liczba tych sieci jest ograniczona do pięciu i dlaczego nie ma wśród nich sieci rombowej oraz sieci kwadratowej centrowanej.

4.3 UKŁADY KRYSTALOGRAFICZNE

W sieci trójwymiarowej komórka elementarna ma postać równoległościanu o krawędziach *a*, *b*, *c* – rys. 4/5. Sieć tę można zbudować przez przesuwanie (translację) tej komórki wzdłuż osi *X*, *Y*, *Z* o wielokrotności odcinków *a*, *b*, *c*. Opis jednej komórki elementarnej wystarcza więc do odtworzenia struktury całej sieci.



4/5
Komórka elementarna sieci trójwymiarowej

Komórka elementarna jest scharakteryzowana przez tzw. *parametry sieci*, tzn. długości jej krawędzi *a*, *b*, *c* oraz kąty α, β, γ zawarte między tymi krawędziami. Wartości tych parametrów wyznaczają rozmiary i kształt komórki.

W zależności od stosunków długości krawędzi oraz wartości zawartych między nimi kątów rozróżnia się 7 układów krystalograficznych, których zasadnicze cechy zestawiono w tabl. 4/1.

W układzie trójskośnym długości wszystkich krawędzi komórki elementarnej są różne i podobnie kąty α, β, γ różnią się między sobą i są różne od 90°. Sieć ta charakteryzuje się najniższym stopniem symetrii.

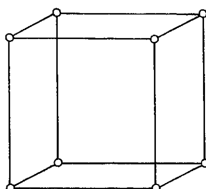
Największą liczbę elementów symetrii wykazuje układ regularny, którego komórka elementarna ma kształt sześcianu. W układzie regularnym występują 3 rodzaje sieci, których komórki elementarne przedstawiono na rys. 4/6. W sieci A_0 krystalizuje np. jedna z odmian alotropowych manganu, w sieci A_1 – złoto, a w sieci A_2 – sód.

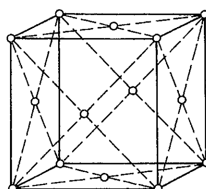
Problem 4.2. Stała sieci kryształu sodu (długość krawędzi komórki sześciennej) wynosi 422,5 pm. Oblicz ile atomów zawiera 1 cm^3 sodu. Wyznacz gęstość tego pierwiastka.

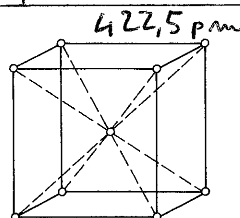
W układzie regularnym krystalizuje ok. 90% zbadanych substancji, w tym znaczna część metali, tlenków, siarczków i halogenków. Kryształy należące do tego układu wykazują izotropię niektórych własności fizycznych, np. rozszerzalności cieplnej, przewodnictwa cieplnego i elektrycznego itd.

4/1 Układy krystalograficzne

Lp.	Nazwa układu	Elementy symetrii	Kształt komórki elementarnej
1	Trójskośny	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90°$	Równoległościan ukośnokątny
2	Jednoskośny	$a \neq b \neq c$ $\alpha = \beta = 90°;\ \gamma \neq 90°$	Graniastosłup o podstawie w kształcie równoległoboku ukośnokątnego
3	Rombowy	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$	Prostopadłościan
4	Tetragonalny	$a = b \neq c$ $\alpha = \beta = \gamma = 90°$	Prostopadłościan o podstawie kwadratowej
5	Regularny	$a = b = c$ $\alpha = \beta = \gamma = 90°$	Sześcian
6	Trygonalny	$a = b = c$ $\alpha = \beta = \gamma \neq 90°$	Romboedr
7	Heksagonalny	$a = b \neq c$ $\alpha = \beta = 90°;\ \gamma = 120°$	Graniastosłup o podstawie rombowej







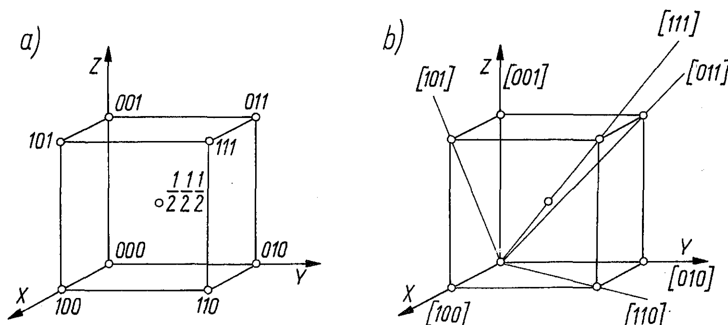
4/6 Komórki elementarne sieci regularnej: A_0 — sieć prosta (niecentrowana); A_1 — sieć płasko centrowana; A_2 — sieć przestrzennie centrowana

4.4 SYMBOLE WĘZŁÓW, PROSTYCH I PŁASZCZYZN SIECIOWYCH

Położenie węzła w trójwymiarowej sieci translacyjnej w stosunku do węzła przyjętego za początek układu współrzędnych jest wyznaczone przez wektor translacji

$$\boldsymbol{R} = m\boldsymbol{a}+n\boldsymbol{b}+p\boldsymbol{c} \qquad (4/2)$$

gdzie $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ – wektory o kierunkach osi krystalograficznych X, Y, Z i długościach równych odpowiednim odcinkom translacji. Położenie węzła jest więc jednoznacznie określone, gdy jest znana uporządkowana trójka liczb m, n, p, którą w związku z tym przyjmuje się w krystalografii za symbol odpowiedniego węzła.



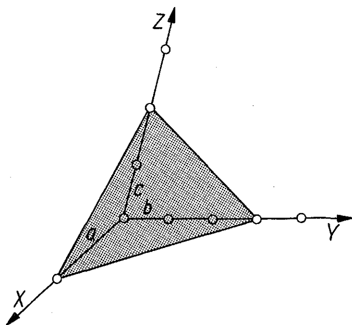
4/7 Symbole węzłów (a) i prostych sieciowych (b) w komórce elementarnej sieci regularnej przestrzennie centrowanej

Przykład oznaczenia węzłów w komórce sieci regularnej przestrzennie centrowanej podano na rys. 4/7a. Dla węzłów w narożach komórki liczby m, n, p są całkowite, natomiast dla węzłów położonych wewnątrz komórki (a w przypadku sieci płasko centrowanej dla węzłów położonych na jej ścianach) przyjmują one wartości ułamkowe.

Wektor wyrażony wzorem (4/2) określa nie tylko położenie węzła, lecz również kierunek prostej sieciowej przechodzącej przez początek układu współrzędnych i przez rozpatrywany węzeł. Tę samą trójkę liczb m, n, p ujętą w nawiasy kwadratowe przyjmuje się zatem za symbol prostej sieciowej lub kierunku. Ze względu na to, że prosta sieciowa przechodzi przez nieskończenie wiele węzłów sieci punktowej, wybór liczb m, n, p nie jest jednoznaczny. W krystalografii przyjmuje się, że liczby te powinny być względem siebie proste (tzn. nie mające wspólnego podzielnika). Przykłady symboli prostych sieciowych pokazano na rys. 4/7b.

Płaszczyzną sieciową nazywamy płaszczyznę poprowadzoną przez 3 węzły sieci translacyjnej nie leżące na prostej. Nachylenie płaszczyzny sieciowej względem osi krystalograficznych określa się za pomocą tzw. *wskaźników Millera*, zdefiniowanych w następujący sposób: Jeżeli rozpatrywana płaszczyzna sieciowa odcina na osiach krystalograficznych X, Y, Z odcinki o długościach ma, nb, pc (rys. 4/8), to wskaźnikami Millera tej płaszczyzny są liczby h, k, l względem siebie proste, spełniające warunki

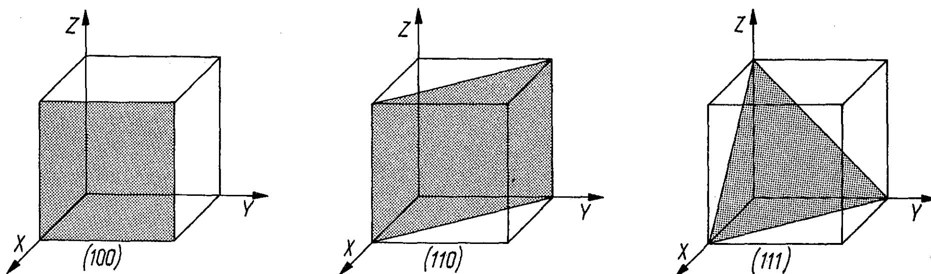
$$h:k:l = \frac{1}{m} : \frac{1}{n} : \frac{1}{p} \tag{4/3}$$



4/8
Płaszczyzna sieciowa w układzie współrzędnych XYZ

Wskaźniki te, ujęte w nawiasy okrągłe, stanowią symbol danej płaszczyzny. Przykładowo, dla płaszczyzny pokazanej na rys. 4/8 mamy: $m = 1$, $n = 3$, $p = 2$, a zatem, jak łatwo stwierdzić, $h = 6$, $k = 2$, $l = 3$. Symbol tej płaszczyzny ma więc postać (623).

Z podanej definicji wynika, że wszystkie płaszczyzny sieciowe równoległe do siebie mają jednakowe wskaźniki Millera. Dla płaszczyzn równoległych do jednej z osi krystalograficznych odpowiedni wskaźnik Millera jest równy zeru (płaszczyzny te przecinają bowiem oś w nieskończoności). Przykłady kilku płaszczyzn sieciowych w komórce sieci regularnej pokazano na rys. 4/9.



4/9 Przykłady symboli płaszczyzn sieciowych w sieci regularnej

4.5 DEFEKTY SIECI KRYSTALICZNEJ

Kryształy rzeczywiste różnią się w istotny sposób od rozpatrywanych w poprzednich punktach kryształów idealnych o doskonale regularnym układzie atomów w przestrzeni. Jednym z czynników zakłócających uporządkowanie atomów w krysztale są drgania cieplne atomów wokół ich położeń równowagi. Ten rodzaj niedoskonałości sieci krystalicznej, należący do kategorii tzw. *zakłóceń energetycznych*, będzie rozpatrywany w następnym punkcie. Obecnie natomiast zajmiemy się tzw. *defektami strukturalnymi*, czyli zakłóceniami spowodowanymi nieprawidłowym rozmieszczeniem niektórych atomów w sieci krystalicznej.

Spośród różnych możliwych rodzajów defektów strukturalnych wymienimy tu następujące rodzaje, które w elementach półprzewodnikowych mają największe znaczenie:

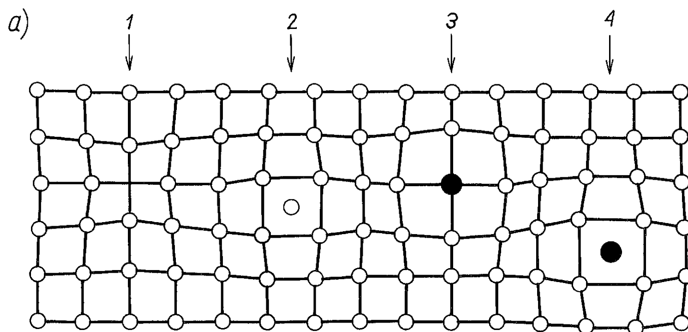
a) *luki węzłowe* (zwane także *wakansami*), czyli węzły sieci krystalicznej przypadkowo nie obsadzone przez właściwe atomy;

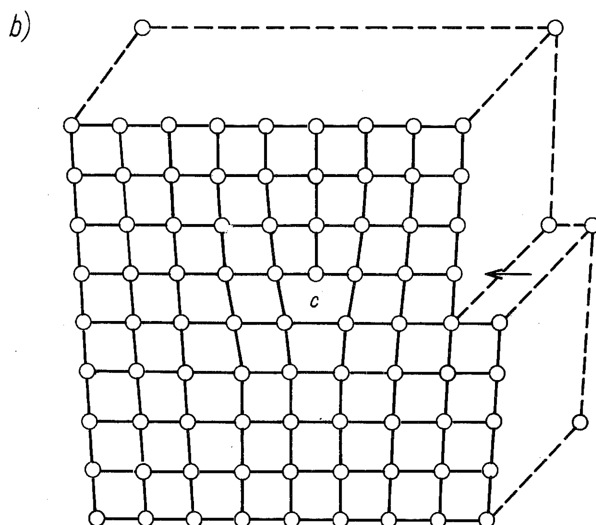
b) *atomy międzywęzłowe*, czyli dodatkowe atomy własne znajdujące się w położeniach, które w idealnej sieci nie powinny być obsadzone;

c) *obce atomy*, zastępujące w węzłach sieci atomy własne, bądź też zajmujące położenia międzywęzłowe;

d) *dyslokacje*, czyli zaburzenia ciągłości sieci, polegające na przesunięciu części sieci w stosunku do reszty.

Defekty te przedstawiono na rys. 4/10. W górnej części rysunku pokazano tzw. *defekty punktowe*, w dolnej zaś – najprostszy rodzaj dyslokacji – tzw. *dyslokację krawędziową*. Polega ona na tym, że w pewnej płaszczyźnie sieciowej





4/10 Defekty sieci krystalicznej: a) defekty punktowe; b) dyslokacja krawędziowa (strzałka wskazuje kierunek, w którym górna część sieci krystalicznej została przesunięta w stosunku do dolnej)
1 — luka węzłowa; *2* — atom własny w położeniu międzywęzłowym; *3* — atom obcy w położeniu węzłowym; *4* — atom obcy w położeniu międzywęzłowym

ciągłość sieci urywa się wzdłuż linii (zwanej *krawędzią*), oznaczonej na rysunku symbolem c.

Bardzo ważnym defektem strukturalnym jest powierzchnia kryształu, na której raptownie urywa się periodyczna struktura sieci krystalicznej. Ponadto, powierzchnia kryształu jest narażona na oddziaływanie otoczenia zewnętrznego. Na powierzchni gromadzą się zatem (zostają zaadsorbowane) znaczne ilości obcych atomów.

Defekty strukturalne powstają zarówno w procesie hodowania kryształu, jak i po zakończeniu tego procesu. Są one przy tym niemożliwe do uniknięcia. Gdyby nawet wyobrazić sobie idealnie czysty kryształ (tzn. kryształ nie zawierający obcych atomów), to w temperaturze różnej od zera bezwzględnego w warunkach równowagi termodynamicznej będzie on zawierał ściśle określoną liczbę luk węzłowych i atomów międzywęzłowych. Wynika to z tego, że wskutek drgań

cieplnych atomów istnieje określony stopień nieuporządkowania struktury krystalicznej. Jest on tym większy, im wyższa jest temperatura kryształu. Działania zewnętrznych czynników fizycznych (zwłaszcza zaś działania mechaniczne i termiczne, oraz niektóre rodzaje promieniowania) powodują powstawanie dodatkowych defektów, a także przemieszczanie się defektów w krysztale.

Kryształy rzeczywiste zawierają zawsze wtrącenia chemiczne w postaci pewnej liczby obcych atomów. Kryształy naturalne uważa się za bardzo czyste, jeżeli zawartość atomów własnych wynosi co najmniej 99,99%, tzn. jeżeli nie więcej niż 1 obcy atom przypada na 10^4 atomów własnych.

W elektronice półprzewodnikowej czystość taka byłaby absolutnie niewystarczająca. Rozwój tej dziedziny stał się możliwy dzięki opracowaniu takich metod oczyszczania materiałów półprzewodnikowych oraz metod hodowania kryształów[1)], przy których liczbę niepożądanych obcych atomów (czyli tzw. *zanieczyszczeń*) udaje się utrzymać w granicach 1 atomu na 10^7–10^9 atomów własnych. W najczyściejszych kryształach 1 obcy atom przypada więc na miliard atomów własnych[2)].

Defekty sieci krystalicznej, a szczególnie obce atomy wpływają w zasadniczy sposób na własności fizyczne (zwłaszcza na własności elektryczne i optyczne) ciał stałych. Najbardziej czułe są pod tym względem półprzewodniki, w których wprowadzenie nawet bardzo małych ilości atomów pewnych pierwiastków (np. 1 atom na miliard atomów własnych) może spowodować wzrost przewodności elektrycznej o kilka rzędów wielkości.

W celu osiągnięcia określonych własności fizycznych półprzewodników stosuje się zatem *domieszkowanie* kryształów. Proces ten polega na *celowym* wprowadzeniu atomów obcych pierwiastków w czasie hodowania kryształu lub po jego wyhodowaniu. W pierwszym przypadku atomy domieszkowe wbudowują się w sieć krystaliczną w czasie jej powstawania, w drugim zaś – są one wprowadzane do wyhodowanego kryształu z zewnątrz. Jedna z najczęściej stosowanych metod polega na tym, że atomy domieszkowe osadzają się na powierzchni kryształu, skąd następnie wnikają do jego wnętrza wskutek dyfuzji[3)].

Domieszkowanie jest podstawową metodą wytwarzania elementów półprzewodnikowych i układów scalonych.

[1)] Najszerzej stosowaną w świecie metodą hodowania monokryształów półprzewodników jest tzw. *metoda Czochralskiego*, opracowana w 1917 r. przez Jana Czochralskiego (1885-1953), późniejszego profesora Politechniki Warszawskiej.

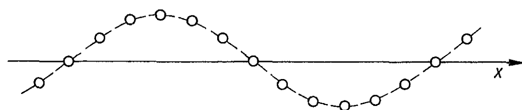
[2)] Istotna jest przy tym nie tylko ilość, lecz również rodzaj obcych atomów. Przykładowo, w krzemie przeznaczonym do produkcji elementów półprzewodnikowych i układów scalonych na 10^9 atomów własnych dopuszcza się co najwyżej 300 atomów węgla, 50 atomów tlenu, ale tylko 1 atom pierwiastków III lub V grupy układu okresowego i 0,1 atomu z grupy metali ciężkich.

[3)] *Dyfuzja* jest to samorzutny proces prowadzący do wyrównywania się rozkładu koncentracji cząstek materialnych wskutek ich nieuporządkowanych ruchów cieplnych. Prowadzi on m.in. do mieszania się gazów oraz stykających się ze sobą cieczy i ciał stałych.

4.6 DRGANIA SIECI KRYSTALICZNEJ

Atomy tworzące sieć krystaliczną ciała stałego wykonują nieustannie chaotyczne drgania wokół swych położeń równowagi. Drgania każdego atomu naruszają równowagę sąsiednich atomów, w wyniku czego rozciągają się one na całą sieć w postaci fal sprężystych rozchodzących się w krysztale we wszystkich możliwych kierunkach. Fale te nazywamy *falami sieciowymi*. Każdej składowej takiej fali można przypisać określoną pulsację Ω i określoną wartość wektora falowego K[1]. Widmo częstotliwości drgań sieci obejmuje zakres od zera do ponad 10^{13} Hz. Fale sieciowe o najmniejszych częstotliwościach są to dobrze znane ze szkoły średniej fale akustyczne.

Zamiast rozpatrywać drgania 3-wymiarowej sieci krystalicznej, posłużymy się prostym 1-wymiarowym modelem w postaci periodycznego łańcucha złożonego z oddziałujących na siebie atomów. Jeden z możliwych sposobów drgań takiego łańcucha przedstawiono na rys. 4/11.

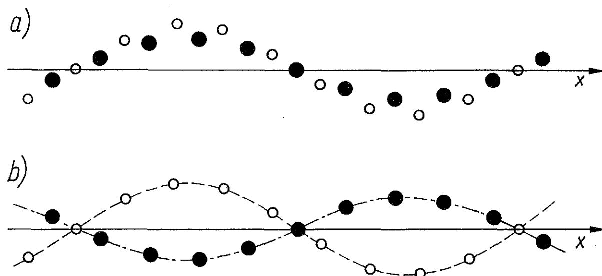


4/11 Drgania 1-wymiarowego modelu sieci krystalicznej zbudowanej z jednakowych atomów

W łańcuchu zbudowanym nie z jednego, lecz z dwóch (lub więcej) rodzajów atomów, mogą występować następujące 2 typy drgań:

a) *drgania akustyczne* (rys. 4/12a), w których oba rodzaje atomów drgają w tych samych fazach, oraz

b) *drgania optyczne* (rys. 4/12b), w których atomy różnych rodzajów drgają w przeciwnych fazach. Drgania tego typu mogą być wywołane zmiennym polem elektrycznym, np. falą elektromagnetyczną o częstotliwości należącej zwykle do zakresu podczerwieni – tym tłumaczy się nazwa „drgania optyczne"[2].



4/12 Drgania 1-wymiarowego modelu sieci krystalicznej zbudowanej z 2 rodzajów atomów: a) drgania akustyczne; b) drgania optyczne

---

[1] W celu odróżnienia tych wielkości od analogicznych wielkości charakteryzujących drgania elektromagnetyczne będziemy je oznaczali dużymi literami.

[2] Drgania optyczne mogą występować również w kryształach zbudowanych z atomów jednego rodzaju. W kryształach tych nie ma jednakże naturalnych dipoli elektrycznych, które mogą być łatwo pobudzone do drgań elektrycznych. Z tego względu oddziaływanie fali elektromagnetycznej na te kryształy jest znacznie słabsze.

Drgania akustyczne obejmują zakres mniejszych częstotliwości (od zera do pewnej częstotliwości granicznej) i mniejszych wartości energii, natomiast drgania optyczne obejmują tylko zakres największych częstotliwości i są związane ze stosunkowo wysokoenergetycznymi zjawiskami elektromagnetycznymi.

Fale sieciowe można opisać za pomocą praw mechaniki kwantowej. Ich energia i wektor falowy są więc skwantowane, podobnie jak to występuje w przypadku fal elektromagnetycznych. Kwantem pola elektromagnetycznego jest, jak wiadomo, foton, którego energia wyraża się wzorem (1/33) lub równoważnym mu wzorem $W = \hbar\omega$, a pęd – wzorem (1/35).

Przez analogię do pojęcia fotonu wprowadza się pojęcie kwantu wzbudzenia pola drgań sieci krystalicznej, który nazywamy *fononem*[1]. Energia fononu wyraża się tym samym wzorem co energia fotonu, a więc wynosi

$$W = \hbar\Omega \tag{4/4}$$

Energie fononów zawierają się w granicach od zera do kilkudziesięciu meV (np. dla krzemu – do 63 meV). Zgodnie z przeprowadzonym tu podziałem typów drgań sieci krystalicznej, rozróżnia się fonony *akustyczne* i *optyczne*.

Fonon ma niektóre cechy takich cząstek, jak foton czy elektron swobodny. Nie jest on jednakże cząstką w pełnym znaczeniu tego słowa, nie ma bowiem samoistnego bytu. Fonon jest tylko pewnym dogodnym sposobem opisu stanu wzbudzenia sieci krystalicznej. Twory tego rodzaju nazywamy *quasi-cząstkami* (czyli niby-cząstkami).

Fonon nie ma również określonego pędu. W oddziaływaniach z innymi cząstkami zachowuje się on jednakże tak, jak gdyby miał pęd

$$p = \hbar K = \frac{h}{\lambda} \tag{4/5}$$

gdzie λ – długość fali sieciowej. Wielkość wyrażoną wzorem (4/5) nazywamy, ściśle biorąc, *quasi-pędem fononu*.

Stan wzbudzenia sieci krystalicznej można traktować jako gaz złożony z fononów poruszających się swobodnie wewnątrz kryształu. W wyniku oddziaływania sieci z cząstkami (np. elektronami lub fotonami) energia drgań sieci, a wraz z nią liczba fononów, może się powiększać lub zmniejszać. W pierwszym przypadku mówimy o emisji fononów, w drugim zaś – o ich absorpcji.

Istnienie fononów zostało udowodnione doświadczalnie. Przy rozpraszaniu promieni rentgenowskich lub neutronów w kryształach obserwuje się mianowicie zmiany energii i pędu (fotonu wzgl. neutronu), odpowiadające emisji lub absorpcji jednego lub większej liczby fononów. Posługując się pojęciem fononów, można wyjaśnić wiele zagadnień dotyczących oddziaływania atomów kryształu z różnego rodzaju cząstkami i promieniowaniem, np. zjawisko rozpraszania elektronów na drganiach cieplnych sieci w metalach i półprzewodnikach.

---

[1] Nazwa ta pochodzi od greckiego słowa „phone" (głos, dźwięk).

Odchylenia atomów od ich położeń równowagi wywołane drganiami cieplnymi sieci krystalicznej są dość znaczne. W temperaturze pokojowej są one rzędu 10–20 pm, co stanowi kilka procent stałej sieci. Przy ogrzewaniu kryształu intensywność drgań atomów wzrasta (liczba fononów powiększa się), a w temperaturze topnienia staje się tak duża, że następuje zniszczenie uporządkowania atomów, a więc przejście ciała w stan ciekły.

5

WIĄZANIA ATOMÓW W KRYSZTALE

Rozważania zawarte w poprzednim rozdziale dotyczyły struktury geometrycznej ciał krystalicznych. Stwierdziliśmy tam, że atomy w ciele stałym tworzą charakterystyczną dla tego ciała sieć krystaliczną, co uzewnętrznia się tym, że pewne kąty i pewne odległości mają ściśle określone wartości, jednakowe w całym krysztale. Umożliwia to odtworzenie całej sieci krystalicznej na podstawie znajomości jednej komórki elementarnej.

W dotychczasowych rozważaniach nie wnikaliśmy jednakże w przyczyny tej regularności, ani też w naturę sił wiążących atomy w kryształ i zapewniających jego spójność. Siły te mają charakter elektryczny i, podobnie jak w przypadku wiązań chemicznych, powstają głównie dzięki elektronom walencyjnym atomów tworzących kryształ.

Rozdział ten jest poświęcony omówieniu różnych rodzajów wiązań, z jakimi spotykamy się w ciałach krystalicznych. Rodzaj wiązania ma niezwykle ważny wpływ na własności fizyczne tych ciał. Jak zobaczymy, decyduje on m.in. o tym, że jedne ciała (metale) są plastyczne i odznaczają się dobrym przewodnictwem elektrycznym, inne zaś (dielektryki) są kruche i prawie nie przewodzą prądu elektrycznego, wreszcie jeszcze inne (półprzewodniki) mają specyficzne własności i w pewnych warunkach (np. przy oświetleniu promieniowaniem o dostatecznie dużej energii fotonów) mogą przewodzić prąd elektryczny.

5.1 RODZAJE WIĄZAŃ KRYSTALICZNYCH

Siły występujące między atomami tworzącymi sieć krystaliczną wynikają głównie z elektrostatycznego przyciągania między ujemnymi ładunkami elektronów walencyjnych a dodatnimi ładunkami rdzeni atomowych. Powstają one wskutek tego, że rozkład elektronów walencyjnych wokół rdzenia atomowego jest w kry-

sztale inny niż w odizolowanym atomie. W zależności od rozkładu tych elektronów rozróżnia się następujące podstawowe rodzaje wiązań krystalicznych:

a) *wiązania jonowe* – występujące między atomami różniącymi się znacznie tzw. *elektroujemnością*[1];

b) *wiązania kowalencyjne* – występujące między atomami tego samego pierwiastka; wiązanie to jest także nazywane *homopolarnym* (tzn. o takich samych biegunach) w odróżnieniu od wiązania jonowego, które jest nazywane *heteropolarnym* (tzn. o różnych biegunach);

c) *wiązania metaliczne* – występujące w metalach i ich stopach;

d) *wiązania molekularne* – występujące w kryształach, które można traktować jako zbudowane nie z pojedynczych atomów, lecz z cząsteczek chemicznych, zachowujących do pewnego stopnia w krysztale swoje indywidualne cechy. Cząsteczki te są ze sobą związane słabymi siłami, zwanymi *siłami van der Waalsa.* Ten typ wiązania występuje również w kryształach gazów szlachetnych.

Wymienione wyżej rodzaje wiązań rzadko występują w czystej postaci. Najczęściej mają one charakter pośredni (np. w kryształach większości związków chemicznych wiązanie ma charakter pośredni między jonowym a kowalencyjnym) lub mieszany[2]. Istnieją również kryształy, w których występują wiązania nie dające się zaliczyć do żadnej z podanych kategorii (przykładem takiego kryształu jest lód).

O tym, jak istotny wpływ na własności fizyczne kryształu ma rodzaj występujących w nim wiązań, można przekonać się na przykładzie diamentu i grafitu. Oba te kryształy są zbudowane z atomów tego samego pierwiastka (węgla), różnią się jednakże sposobem ułożenia atomów w sieć krystaliczną. W wyniku tego mają one diametralnie różne własności fizyczne. Diament (wiązanie kowalencyjne) jest przezroczysty, ma najwyższą znaną twardość i w normalnych warunkach nie przewodzi prądu elektrycznego, grafit natomiast jest bardzo miękki (twardość 1 w skali Mohsa), jest nieprzezroczysty i odznacza się dobrym przewodnictwem elektrycznym.

Miarą wielkości sił wiążących atomy w kryształ jest *energia wiązania.* Jest to energia potrzebna do rozłożenia kryształu na pojedyncze neutralne atomy.

---

[1] *Elektroujemność* jest to zdolność atomu do przyciągania elektronów w warunkach, gdy atom znajduje się w ścisłym kontakcie z innymi atomami, tworząc np. cząsteczkę chemiczną.

Istnieje kilka umownych skal elektroujemności pierwiastków. Najpowszechniej używana jest skala Paulinga opracowana na podstawie zmiany energii potencjalnej układu w czasie tworzenia związku chemicznego. W obrębie każdego okresu elektroujemność pierwiastków wzrasta w miarę przesuwania się od I do VII grupy układu okresowego. Największą elektroujemność wykazują fluor i tlen, najmniejszą zaś metale alkaliczne. Pierwiastki o bardzo małej elektroujemności nazywa się często *elektrododatnimi.*

[2] Np. kryształy grafitu są zbudowane z płaskich warstw, w których wiązania atomów mają częściowo charakter kowalencyjny, częściowo zaś metaliczny. Warstwy te są ze sobą związane stosunkowo słabymi wiązaniami o charakterze molekularnym. Z tego względu grafit wykazuje silną anizotropię własności mechanicznych, elektrycznych itd.

Wyraża się ją w elektronowoltach na atom lub w dżulach na mol. Energia ta zawiera się w bardzo szerokich granicach od 0,02 eV/atom dla neonu (temperatura topnienia −249°C) do 8,9 eV/atom dla wolframu (temperatura topnienia 3370°C).

5.2 WIĄZANIE JONOWE

Jak już zaznaczono, wiązanie jonowe występuje w połączeniach atomów, różniących się znacznie elektroujemnością, a zatem należących do możliwie oddalonych od siebie grup układu okresowego pierwiastków. Najbardziej typowym przykładem takich połączeń są związki pierwiastków I i VII grupy układu okresowego, a więc halogenki metali alkalicznych.

Wiązanie jonowe rozpatrzymy tu na przykładzie soli kuchennej (NaCl). Atom sodu ma, jak wiemy, całkowicie zapełnione powłoki K i L oraz 1 elektron walencyjny należący do powłoki M. Elektron ten jest stosunkowo słabo związany z rdzeniem atomowym (energia jonizacji atomu sodu wynosi 5,14 eV). Z kolei chlor ma konfigurację elektronową $3s^2 3p^5$, ma więc 7 elektronów walencyjnych i do zapełnienia podpowłoki $3p$ brak mu 1 elektronu. Pierwiastek ten odznacza się zatem znaczną elektroujemnością (powinowactwo elektronowe atomu chloru wynosi 3,71 eV[1)]).

Cząsteczka NaCl powstaje wskutek przeniesienia elektronu walencyjnego z atomu Na do atomu Cl. W wyniku tego atom sodu przekształca się w jon Na^+, a atom chloru – w jon Cl^-. W obu tych jonach rozkład ładunku elektrycznego wykazuje symetrię sferyczną. Jony te można więc w przybliżeniu traktować jako naładowane kule.

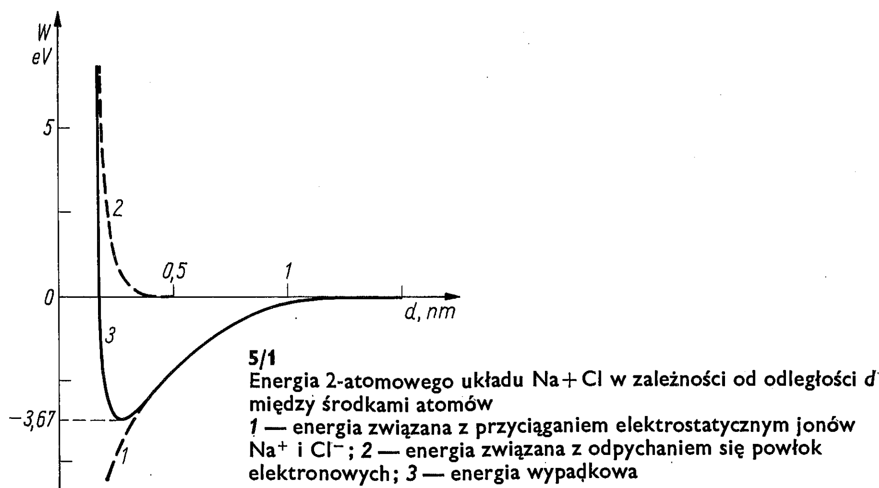
Między jonami występują siły przyciągania elektrostatycznego. W miarę zbliżania się jonów do siebie, pojawiają się jednak również przeciwnie skierowane siły wywołane odpychaniem się powłok elektronowych obu atomów.

Wartość i kierunek siły wypadkowej występującej między atomami Na i Cl w zależności od odległości d między ich środkami można odczytać z pokazanego na rys. 5/1 wykresu przedstawiającego zależność całkowitej energii układu Na+Cl od odległości d. Energia ta jest mierzona w stosunku do stanu, w którym atomy Na i Cl nie są zjonizowane i znajdują się w tak dużej odległości od siebie, że ich wzajemne oddziaływanie może być pominięte. Liniami przerywanymi pokazano na rysunku energie związane z siłami przyciągającymi i z siłami odpychającymi.

Jeżeli wyobrazimy sobie, że zbliżamy do siebie atomy Na i Cl, które znajdowały się np. w odległości 2 nm, to, jak widać na rysunku, w miarę zmniejszania odległości d energia układu początkowo maleje. Świadczy to o przewadze sił przyciągania nad siłami odpychającymi. W pewnej odległości (która dla ato-

---

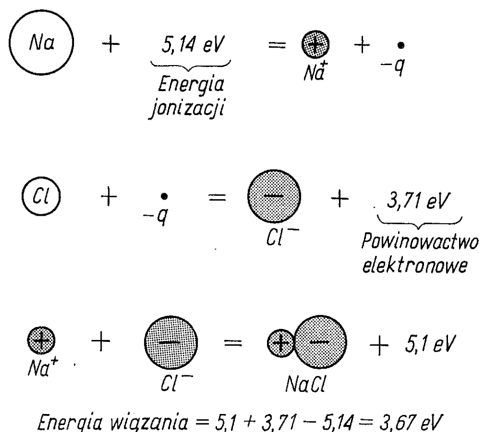
[1)] *Powinowactwo elektronowe* jest to energia wydzielana w procesie przyłączania przez dany atom elektronu znajdującego się początkowo w nieskończenie dużej odległości.



5/1
Energia 2-atomowego układu Na+Cl w zależności od odległości *d* między środkami atomów
1 — energia związana z przyciąganiem elektrostatycznym jonów Na^+ i Cl^-; *2* — energia związana z odpychaniem się powłok elektronowych; *3* — energia wypadkowa

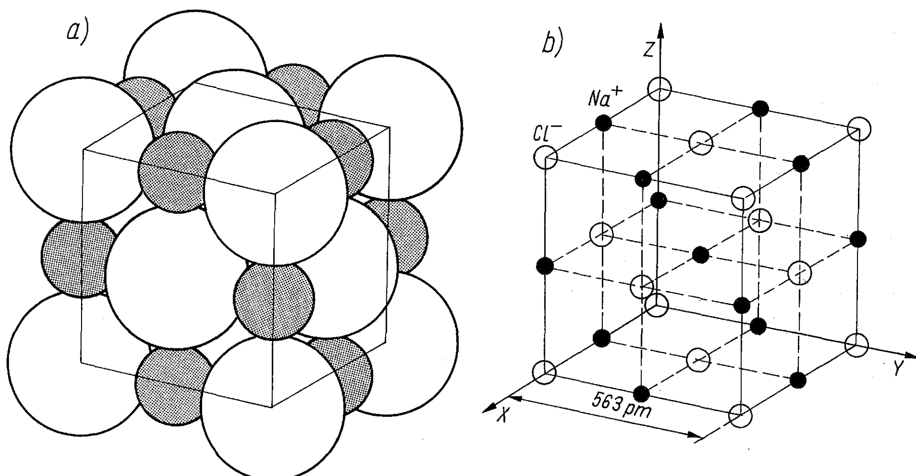
mów Na i Cl wynosi 0,28 nm) energia układu osiąga wartość minimalną, a zatem siły przyciągania równoważą się z siłami odpychającymi. Układ znajduje się wówczas w stanie równowagi, tworząc cząsteczkę NaCl. Ta minimalna wartość energii wynosi −3,67 eV. Jest ona ujemna, ponieważ do rozdzielenia cząsteczki NaCl na elektrycznie obojętne atomy Na i Cl należy dostarczyć energii z zewnątrz.

Wartość energii wiązania cząsteczki NaCl wynika z rachunku przedstawionego na rys. 5/2.



5/2
Wyznaczenie energii wiązania cząsteczki NaCl

Kryształy jonowe, podobnie jak cząsteczki o wiązaniu jonowym, są zbudowane z dodatnich i ujemnych jonów. Jony te są ułożone w taki sposób, że siły przyciągania między jonami różnoimiennymi przewyższają siły odpychania między jonami o jednakowych znakach. Jest to możliwe dzięki temu, że w najbliższym sąsiedztwie danego jonu znajdują się jony przeciwnego znaku. Przykładowo, w krysztale NaCl (rys. 5/3) każdy jon Na^+ jest otoczony przez 6 jonów Cl^-,



5/3 Kryształ NaCl: a) model sporządzony przy zachowaniu właściwych proporcji promieni jonowych w stosunku do stałej sieci; b) model uproszczony
Promienie: jonu Na^+ — 98 pm, jonu Cl^- — 183 pm

a każdy jon Cl^- ma 6 najbliższych sąsiadów w postaci jonów Na^+. W pierwszym przybliżeniu jony te można traktować jako naładowane sztywne kule, stykające się ze sobą w sposób zapewniający najgęściejsze upakowanie. Przez stykanie się jonów przeciwnego znaku należy rozumieć, że znajdują się one w takiej odległości od siebie, iż siły przyciągania równoważą się z siłami odpychania powłok elektronowych.

Na podstawie rys. 5/3 można stwierdzić, że kryształy soli kuchennej tworzą sieć regularną płasko centrowaną. Bazę tej sieci stanowi para atomów Na i Cl. Nie należy jednak sądzić, że pary te zachowują indywidualne cechy cząsteczki. W kryształach jonowych nie można wyodrębnić takich grup atomów, które odpowiadałyby cząsteczkom. Oznacza to, że każdy jon jest jednakowo silnie związany ze wszystkimi swoimi najbliższymi sąsiadami. W wyniku tego cały kryształ jonowy można traktować jako jedną gigantyczną cząsteczkę.

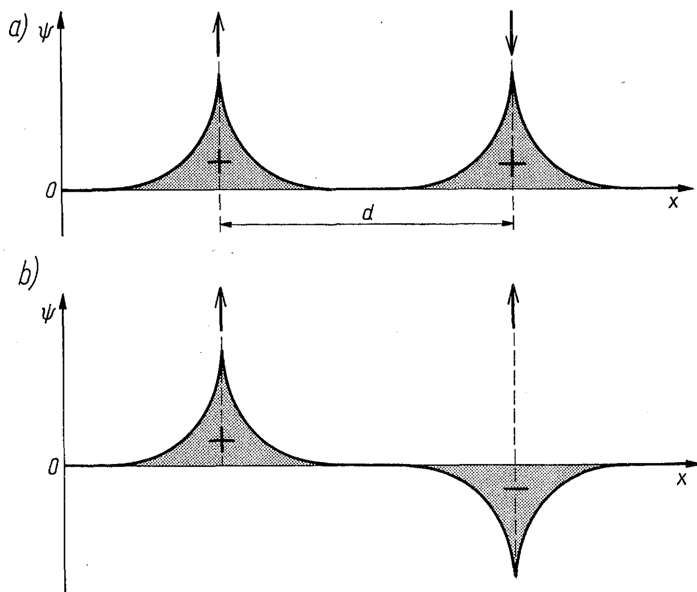
W kryształach, w których wiązanie ma silnie jonowy charakter, elektrony są mocno związane z macierzystymi jonami, w związku z czym nie mogą one przemieszczać się pod wpływem przyłożonego pola elektrycznego. Kryształy te nie wykazują więc przewodnictwa *elektronowego*. Są one zatem zaliczane do klasy *dielektryków*[1)].

[1)] Kryształy jonowe mogą w podwyższonych temperaturach wykazywać przewodnictwo *jonowe*. Przyczyną występowania swobodnych jonów w tych kryształach są lokalne niedoskonałości sieci krystalicznej. Przewodnictwo jonowe charakteryzuje się tym, że w przeciwieństwie do przewodnictwa elektronowego, z przepływem prądu elektrycznego związany jest transport materii.

5.3 WIĄZANIE KOWALENCYJNE

Ten rodzaj wiązania występuje w czystej postaci w substancjach zbudowanych z jednakowych atomów. Własności wiązania kowalencyjnego poznamy na przykładzie najprostszego układu utworzonego z jednakowych atomów, tj. cząsteczki H_2.

Wyobraźmy sobie w tym celu, że zbliżamy do siebie 2 atomy wodoru, z których każdy znajduje się w stanie podstawowym, tzn. w stanie 1*s*. Przy dostatecznie dużej odległości *d* między protonami rozwiązanie równania Schrödingera dla każdego z tych atomów jest prawie dokładnie takie, jak dla atomu odizolowanego. Funkcje falowe rozpatrywanych atomów mają więc postać pokazaną na rys. 5/4a lub 5/4b. W obu przypadkach są one identyczne co do wartości bezwzględnej, mogą natomiast mieć ten sam znak lub też różnić się znakiem.



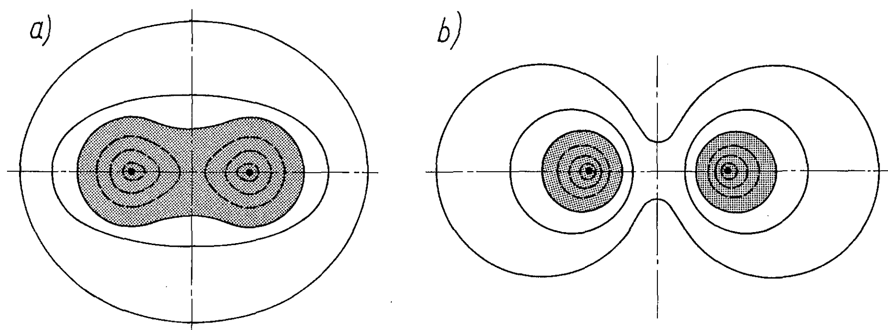
5/4 Funkcje falowe 2 atomów wodoru znajdujących się w odległości *d* znacznie większej niż a_0: a) przypadek symetryczny; b) przypadek antysymetryczny

Znak funkcji falowej dla pojedynczego atomu nie ma żadnego znaczenia, bowiem sens fizyczny ma jedynie kwadrat modułu funkcji falowej. W układzie dwuatomowym zgodność lub niezgodność znaków funkcji falowych ma natomiast zasadnicze znaczenie. Z praw symetrii funkcji falowych, których nie będziemy tu szerzej omawiali, wynika bowiem, że w układzie tym są możliwe następujące dwie sytuacje: albo funkcje falowe mają ten sam znak (przypadek ten określamy jako *symetryczny*), a wówczas spiny elektronów są skierowane antyrównolegle, albo też funkcje falowe różnią się znakiem (przypadek *antysymetryczny*), a spiny są skierowane równolegle. Orientacje spinów zaznaczono na rys. 5/4 strzałkami.

W warunkach przedstawionych na rys. 5/4 orbitale obu atomów prawie nie zachodzą na siebie, co oznacza, że atomy te nie oddziałują ze sobą. Przy zbliżaniu atomów następuje częściowe *nakrywanie* się chmur elektronowych. Nie sprowadza się ono jednakże do zwykłego sumowania funkcji falowych, towarzyszy mu bowiem przegrupowanie rozkładu ładunku (tzn. zmiana kształtu chmur elektronowych) i zmiana energii układu.

Przy dostatecznie małej odległości między protonami mówienie o przynależności danego elektronu do określonego atomu nie ma sensu. W układzie powstaje nowy stan kwantowy, polegający na tym, że oba elektrony należą jednocześnie do obu atomów. Mówimy wówczas, że zostały one *skolektywizowane*, czyli *uwspólnione*. Funkcję falową opisującą gęstość utworzonej w ten sposób chmury elektronowej nazywamy *orbitalem molekularnym*.

Rozkład gęstości ładunku elektronowego dla obu rozpatrywanych wyżej przypadków przedstawiono na rys. 5/5. Linie pokazane na tym rysunku są to warstwice powierzchni przedstawiającej rozkład gęstości chmury elektronowej w płaszczyźnie przechodzącej przez oba protony.



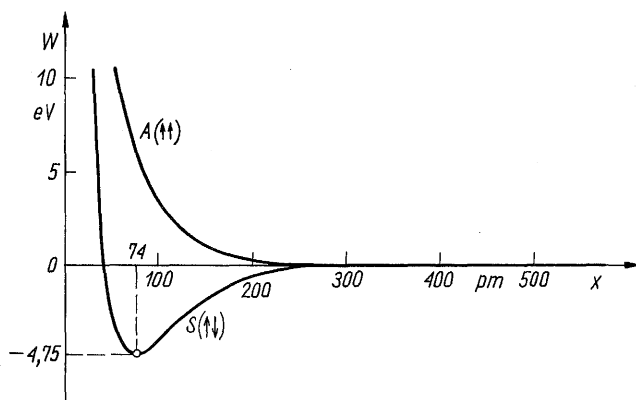
5/5 Rozkład gęstości chmury elektronowej w układzie złożonym z 2 atomów wodoru: a) przypadek symetryczny (orbital wiążący) — atomy tworzą cząsteczkę H_2; b) przypadek niesymetryczny (orbital antywiążący). W celu zwiększenia czytelności rysunku zacieniono obszary, w których gęstość ładunku elektronowego przekracza pewną wartość jednakową na obu wykresach

W przypadku symetrycznym (spiny elektronów ustawione antyrównolegle) następuje podczas kolektywizacji elektronów częściowe wciągnięcie chmur elektronowych z obszarów peryferyjnych do obszaru między protonami. Prawdopodobieństwo znalezienia elektronu w obszarze między protonami ulega więc zwiększeniu, co prowadzi do powstania sił elektrostatycznych, które starają się zbliżyć protony do siebie. Orbital taki nazywamy więc *orbitalem wiążącym*. Przy pewnej odległości między protonami siły przyciągania wywierane na nie przez elektrony walencyjne równoważą się z siłami wzajemnego odpychania jednoimiennych ładunków tych protonów. Układ znajduje się wówczas w stanie równowagi, tworząc dwuatomową cząsteczkę wodoru.

Z kolei w przypadku antysymetrycznym (spiny ustawione równolegle) gęstość chmury elektronowej w obszarze między protonami jest mała. Wynika to

z zakazu Pauliego, zgodnie z którym 2 elektrony nie mogą znajdować się w tym samym stanie kwantowym (w tym przypadku oba elektrony znajdują się w stanie 1*s*, a więc mają jednakowe liczby kwantowe *n*, *l*, *m*, a ponieważ ich spiny są równoległe, zatem również czwarte liczby kwantowe są jednakowe). Ładunek elektronowy jest teraz w znacznym stopniu skupiony w obszarach znajdujących się poza protonami, wskutek czego do sił wzajemnego odpychania się protonów dochodzą dodatkowe siły odciągające je od siebie. Układ taki nie może więc znaleźć się w stanie równowagi, a przypisany temu stanowi orbital molekularny nazywamy ***orbitalem antywiążącym.***

W wyniku zbliżania protonów następuje, jak to już zaznaczono, nie tylko zmiana rozkładu gęstości chmury elektronowej, lecz również zmiana całkowitej energii układu. Zależność tej energii od odległości *d* między protonami przedstawiono na rys. 5/6. Na osi rzędnych odłożono tu energię mierzoną w stosunku do stanu, w którym atomy znajdują się nieskończenie daleko od siebie.



5/6 Zależność energii układu złożonego z 2 atomów wodoru od odległości *d* między protonami: S — przypadek symetryczny; A — przypadek antysymetryczny

Górna krzywa na rysunku odpowiada przypadkowi spinów równoległych. Jak widać, w miarę zbliżania protonów energia układu rośnie monotonicznie. Zbliżanie atomów wymaga więc wykonania pracy (tzn. dostarczenia energii z zewnątrz) w celu pokonania sił wzajemnego odpychania się atomów.

Przy antyrównoległym ustawieniu spinów atomy początkowo przyciągają się, a zatem energia układu maleje. Przy odległości równej 1,4 a_0 osiąga ona wartość minimalną, a przy dalszym zmniejszaniu odległości *d* wzrasta wskutek przewagi sił odpychania się protonów nad przyciągającym działaniem elektronów.

Odległość między jądrami sąsiednich atomów w cząsteczce (lub w krysztale) nazywamy *długością wiązania*. Dla cząsteczki H_2 długość ta wynosi 1,4 a_0 = 74 pm. Z rysunku 5/6 można odczytać, że energia wiązania tej cząsteczki wynosi 4,75 eV.

Z podanych tu rozważań wynika, że wiązanie kowalencyjne między 2 atomami powstaje w wyniku utworzenia się wspólnej pary elektronów walencyjnych (po 1 elektronie z każdego atomu) o przeciwnie skierowanych spinach. O takich

elektronach mówimy, że są *sparowane*, lub też, że mają sparowane spiny. W tworzeniu się wiązania kowalencyjnego istotne jest to, że uwspólniając swoje elektrony, atomy wytwarzają stan, którego całkowita energia jest mniejsza niż suma ich energii w warunkach, gdy są one rozdzielone.

5.4 KRYSZTAŁY O WIĄZANIU KOWALENCYJNYM

Typowymi kryształami o wiązaniu kowalencyjnym są kryształy pierwszych 3 pierwiastków należących do IV grupy układu okresowego, tzn. węgla, krzemu i germanu[1]. Dwa ostatnie, a zwłaszcza krzem, odgrywają bardzo ważną rolę w elektronice półprzewodnikowej, dlatego też będziemy im w tej książce poświęcali szczególną uwagę.

Wszystkie wymienione tu pierwiastki krystalizują w układzie, który określamy jako *strukturę diamentu.* Z kryształami o wiązaniu kowalencyjnym zapoznamy się więc na przykładzie tej właśnie odmiany węgla[2].

Atom węgla w stanie podstawowym ma konfigurację elektronową $2s^2 2p^2$ (zob. rys. 3/10). Ma on więc 2 niesparowane elektrony i na tej podstawie można by oczekiwać, że pierwiastek ten powinien być 2-wartościowy[3]. W większości przypadków węgiel jest jednakże pierwiastkiem 4-wartościowym. Jest to wynikiem przeniesienia (czyli tzw. *promocji*) jednego z elektronów podpowłoki $2s$ na jeden z nieobsadzonych orbitali podpowłoki $2p$. W wyniku tego konfiguracja elektronowa przyjmuje postać pokazaną na rys. 5/7a. Jest to tzw. *stan walencyjny* atomu węgla, w którym ma on 4 niesparowane elektrony o równolegle skierowanych spinach.

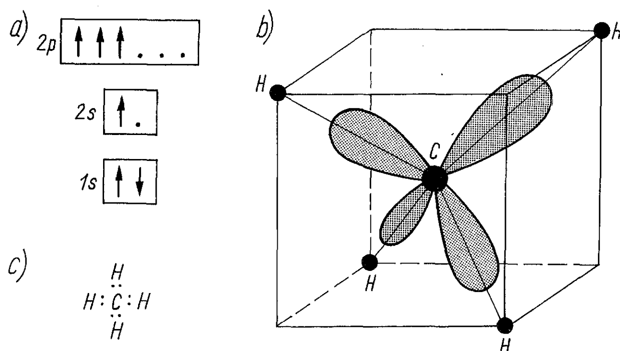
Obecność 4 niesparowanych elektronów umożliwia atomowi węgla przyłączenie np. 4 atomów wodoru, w wyniku czego powstaje cząsteczka metanu CH_4.

W cząsteczce CH_4 orbitale $2s$ i $2p$ atomu węgla nie zachowują swego normalnego kształtu, lecz ulegają zmieszaniu (czyli tzw. *hybrydyzacji*), w wyniku czego powstaje orbital będący kombinacją liniową 1 funkcji falowej $2s$ oraz 3 funkcji falowych $2p$. Nosi on nazwę orbitalu zhybrydyzowanego sp^3. Jego kształt przedstawiono w sposób poglądowy na rys. 5/7b. Chmura elektronowa otaczająca rdzeń atomowy węgla ma jak gdyby 4 równoważne „ramiona" skierowane wzdłuż osi łączących środek czworościanu foremnego (*tetraedru*) z jego narożami. Każde z tych „ramion" tworzy wiązanie z jednym atomem wodoru. Wiązanie to, podobnie jak w cząsteczce H_2, można uważać za wynik oddziaływania zachodzącego między parą elektronów o antyrównoległych spinach. Parę

[1] Dalsze 2 pierwiastki tej grupy to cyna i ołów. Cyna ma dwie odmiany alotropowe; powszechnie znaną odmianą jest tzw. cyna biała, która, podobnie jak ołów, jest metalem. Cyna szara jest półprzewodnikiem o wiązaniu kowalencyjnym.

[2] Jak już zaznaczono, węgiel ma 2 odmiany alotropowe: diament i grafit. W punkcie tym, mówiąc o kryształach węgla, będziemy mieli jednakże na myśli jedynie diament.

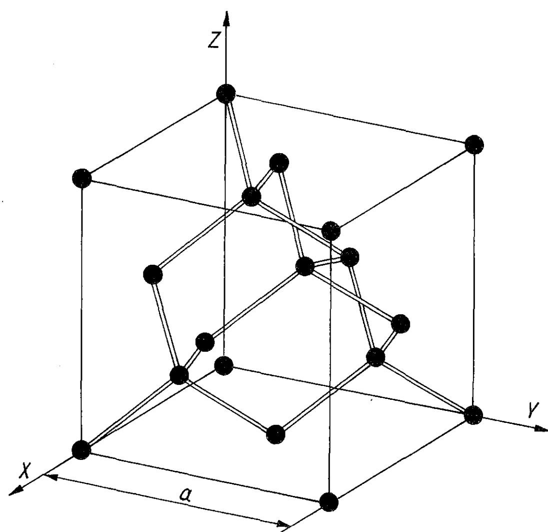
[3] Istotnie, w niektórych związkach chemicznych (np. w tlenku węgla CO) węgiel wykazuje taką wartościowość.



5/7 Stan walencyjny atomu węgla: a) konfiguracja elektronowa; b) orbital atomowy w cząsteczce CH_4; c) wzór strukturalny cząsteczki CH_4

tę stanowi jeden z czterech elektronów walencyjnych atomu węgla oraz elektron atomu wodoru[1]. We wzorach chemicznych oznacza się niekiedy taką parę za pomocą 2 kropek. Przy zastosowaniu tej konwencji otrzymuje się pokazany na rys. 5/7c wzór strukturalny metanu.

O wiązaniu występującym w cząsteczce CH_4 mówimy, że wykazuje *symetrię tetraedryczną*. Taki sam tetraedryczny charakter mają wiązania w krysztale diamentu. Każdy atom węgla ma tu 4 najbliższych sąsiadów, z którymi jest połączony wiązaniami kowalencyjnymi za pomocą uwspólnionej pary elektronów walencyjnych – zob. rys. 5/8.



Stała sieci a w pm	
C	356,7
Si	543,1
Ge	565,8
α-Sn	648,9

5/8 Komórka elementarna sieci diamentu

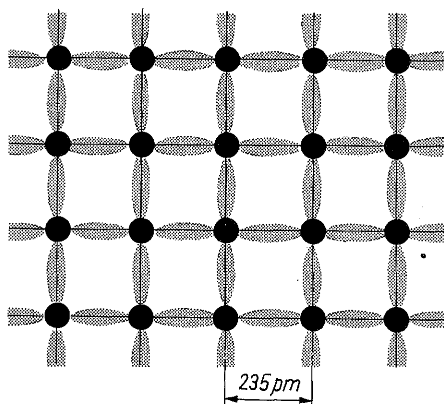
[1] Opis, w którym traktuje się, że każde z 4 „ramion" orbitalu sp^3 reprezentuje jeden elektron walencyjny atomu węgla, jest bardzo obrazowy, ale nieścisły. W rzeczywistości każdy z elektronów walencyjnych atomu węgla ma jednakowe prawdopodobieństwo znalezienia się w którymkolwiek „ramieniu". Ładunek 4 elektronów walencyjnych jest więc rozłożony równomiernie na wszystkie „ramiona" orbitalu.

Wiązanie kowalencyjne ma charakter kierunkowy. Powstaje ono wzdłuż osi, wokół których skupia się największa część chmury elektronowej. W wyniku tego w kryształach o wiązaniu kowalencyjnym atomy ustawiają się nie tylko w określonych odległościach od siebie, lecz także tworzą określoną strukturę przestrzenną o wyróżnionych kierunkach wiązań.

Przedstawioną na rys. 5/8 strukturę diamentu można traktować jako składającą się z 2 sieci regularnych płasko centrowanych, przesuniętych względem siebie o 1/4 głównej przekątnej sześcianu. Inaczej mówiąc, sieć translacyjna diamentu jest siecią regularną typu A_1, a bazę struktury diamentu stanowią 2 jednakowe atomy, znajdujące się w pozycjach 000 i $\frac{1}{4}\frac{1}{4}\frac{1}{4}$.

Problem 5.1. Znając stałą sieci krzemu $a = 543{,}1$ pm, wyznacz długość wiązania (tzn. odległość między środkami sąsiednich atomów) oraz koncentrację atomów (tzn. liczbę atomów w jednostce objętości) w krysztale krzemu.

Przy rozpatrywaniu struktury diamentu często dogodnie jest posługiwać się przedstawionym na rys. 5/9 płaskim modelem, zachowującym podstawową cechę rzeczywistej struktury przestrzennej, polegającą na tym, że każdy atom jest połączony wiązaniami ze swymi 4 najbliższymi sąsiadami. Kółka na tym rysunku reprezentują czterododatnie rdzenie atomowe, złożone w przypadku krzemu ($Z = 14$) z jądra oraz 10 elektronów całkowicie zapełnionych powłok wewnętrznych. Między rdzeniami uwidoczniono uwspólnioną (dla całego kryształu!) chmurę elektronową utworzoną ze wszystkich elektronów walencyjnych. Chmura ta wykazuje zagęszczenia w obszarach między rdzeniami atomowymi[1].



5/9
Uproszczony (dwuwymiarowy) model sieci krystalicznej krzemu (na rysunku zachowano właściwe proporcje promieni jonów Si^{4+} w stosunku do odległości między nimi)

[1] Istnienie takich zagęszczeń można stwierdzić doświadczalnie, badając dyfrakcję promieni rentgenowskich w kryształach.

Kryształy o wiązaniu kowalencyjnym są w dostatecznie niskich temperaturach izolatorami. Brak w nich bowiem swobodnych elektronów, które mogłyby wytwarzać prąd elektryczny pod wpływem doprowadzonego napięcia. Elektrony powłok wewnętrznych są silnie związane z jądrami atomowymi, natomiast elektrony walencyjne tworzą wiązania międzyatomowe i także nie wnoszą przyczynku do prądu elektrycznego. W miarę powiększania temperatury wzrasta energia drgań cieplnych sieci krystalicznej, w wyniku czego część wiązań międzyatomowych ulega zerwaniu. Oswobodzone w ten sposób elektrony mogą się teraz swobodnie poruszać w całym krysztale, a zatem kryształ zaczyna przewodzić prąd elektryczny. Przewodnictwo elektryczne może pojawić się również pod wpływem innych czynników, np. w wyniku absorpcji promieniowania o dostatecznie dużej energii fotonów. Wskutek absorpcji fotonu następuje wyswobodzenie jednego elektronu z wiązań międzyatomowych, co umożliwia przepływ prądu elektrycznego przez kryształ.

Materiały wykazujące opisane wyżej własności noszą nazwę *półprzewodników*. Będą one dokładniej rozpatrywane w następnych dwóch rozdziałach.

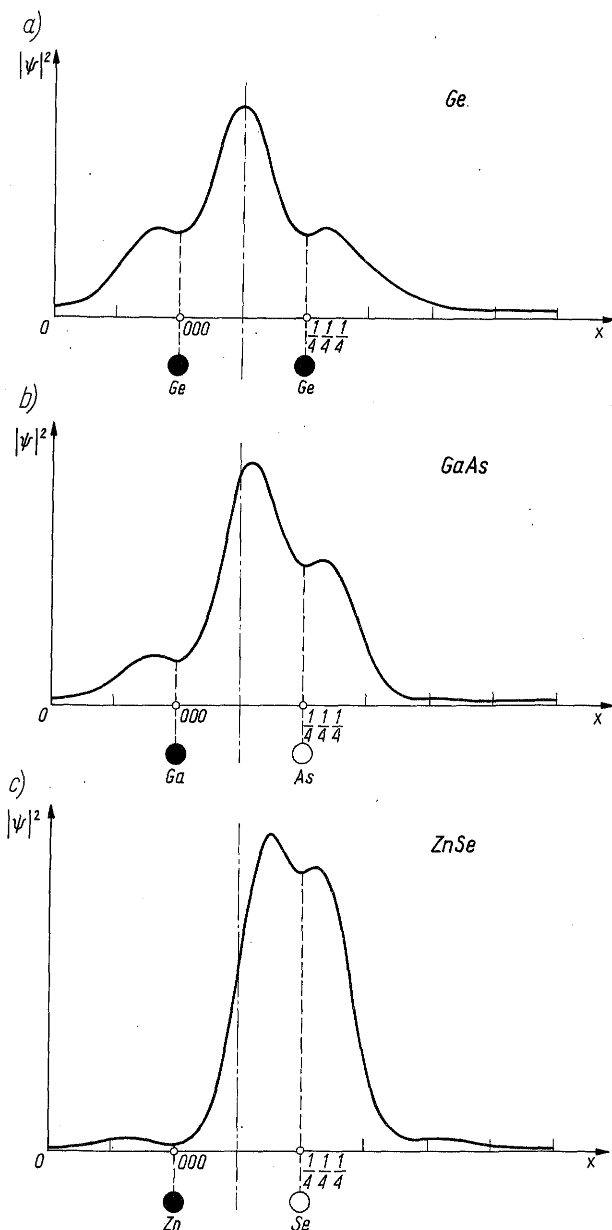
Wiązanie kowalencyjne odznacza się dużą energią wiązania. Jest ona porównywalna z energią wiązania jonowego, pomimo tego, że mamy tu do czynienia z kryształem zbudowanym z elektrycznie obojętnych atomów. Przykładowo dla krzemu energia wiązania wynosi 4,8 eV na atom. Dzięki dużej wartości energii wiązania kryształy o wiązaniu kowalencyjnym charakteryzują się dużą twardością i wysoką temperaturą topnienia, a kierunkowość wiązania sprawia, że są one kruche.

5.5* WIĄZANIE SPOLARYZOWANE

Opisane w poprzednich punktach wiązania jonowe i kowalencyjne są skrajnymi przypadkami, z którymi w rzeczywistości spotykamy się dość rzadko. W większości znanych związków chemicznych sytuacja nie jest zwykle tak krańcowa, tzn. wiązanie ma charakter pośredni między jonowym a kowalencyjnym. Wiązanie takie nazywamy *spolaryzowanym*, przy czym stopień polaryzacji wyrażamy za pomocą tzw. *współczynnika jonowości*[1].

Dla kryształów o czystym wiązaniu kowalencyjnym (np. dla diamentu, krzemu i germanu) rozkład gęstości chmury elektronowej jest symetryczny względem płaszczyzny przechodzącej przez środek odcinka łączącego sąsiednie atomy, a maksimum gęstości ładunku elektronowego przypada w połowie tego odcinka – zob. rys. 5/10a. Wiązanie nie jest więc spolaryzowane, a współczynnik jonowości jest równy zeru. Dla kryształów utworzonych z atomów różnych pierwiastków maksimum gęstości chmury elektronowej jest przesunięte w kie-

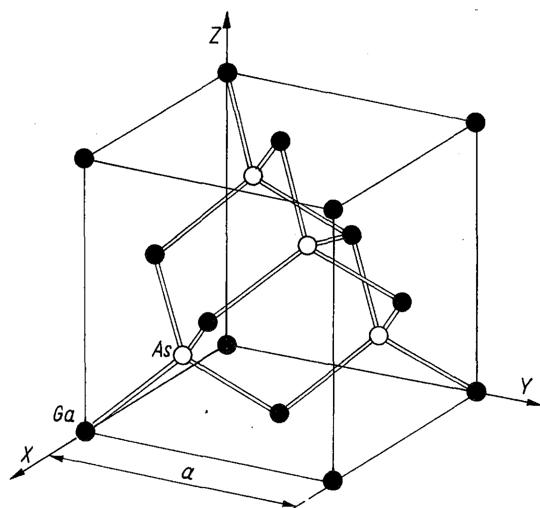
[1] Istnieje kilka sposobów definiowania współczynnika jonowości wiązania. Tu będziemy posługiwali się współczynnikiem wprowadzonym przez Phillipsa (J. C. Phillips, J. A. Van Vechten: Dielectric classification of crystal structures, ionization potentials, and band structures. *Phys. Rev. Letters* **22**, 705 (1969)).



5/10 Rozkład gęstości chmury elektronowej wzdłuż prostej sieciowej [111] w krysztale germanu (a), arsenku galu (b) i selenku cynku (c)
Opracowano na podstawie: J. C. Phillips: *Bonds and bands in semiconductors*. Academic Press, New York 1973, s. 143.

runku atomu pierwiastka bardziej elektroujemnego, a współczynnik jonowości przyjmuje wartość większą od zera.

Jako przykład takiego kryształu rozpatrzmy arsenek galu GaAs, który obok krzemu odgrywa w elektronice półprzewodnikowej bardzo ważną rolę. Na rysunku 5/11 przedstawiono komórkę elementarną tego kryształu. Pod względem geometrycznym przypomina ona komórkę diamentu (rys. 5/8). Podobnie jak w przypadku diamentu można sieć GaAs traktować jako złożoną z 2 podsieci typu A_1 przesuniętych względem siebie o 1/4 głównej przekątnej sześcianu. W arsenku galu jedna podsieć jest utworzona z atomów Ga, druga zaś z atomów As. Strukturę taką nazywamy *strukturą blendy cynkowej*[1)].



	Stała sieci *a* w pm
GaAs	565,3
GaP	545,1
GaSb	609,6
InAs	605,8
InP	586,9
InSb	647,9
CdS	581,8
CdSe	605,2
CdTe	648,1
ZnS	541,0
ZnSe	566,8

5/11 Komórka elementarna arsenku galu i innych związków, krystalizujących w strukturze blendy cynkowej

Z rysunku 5/11 wynika, że każdy atom Ga ma 4 najbliższych sąsiadów w postaci atomów As i odwrotnie, każdy atom As jest związany z 4 atomami Ga. Podobnie jak w strukturze diamentu wiązania te są utworzone przez pary elektronów o przeciwnie skierowanych spinach, z tym jednak, że arsen jako pierwiastek 5-wartościowy dostarcza do tych wiązań 5 swoich elektronów walencyjnych, a gal jako pierwiastek 3-wartościowy – 3 elektrony. Polaryzacja wiązania występującego w GaAs polega na tym, że prawdopodobieństwo znalezienia jednego z tych 8 elektronów w pobliżu atomu As jest większe niż 5/8, a w pobliżu atomu Ga – mniejsze niż 3/8. Jest to spowodowane różnicą elektroujemności tych pierwiastków. Arsen jako atom bardziej elektroujemny przyciąga elektrony silniej niż gal, wskutek czego przekształca się on w ujemnie naładowany jon, podczas gdy gal staje się jonem dodatnim. Zjawisko to jest analogiczne z opisanym w p. 5.2 przekształceniem atomów Na i Cl w jony Na^+ i Cl^-, z tym tylko

[1)] Nazwa ta pochodzi od minerału, będącego bardzo rozpowszechnioną postacią krystaliczną siarczku cynku ZnS.

zastrzeżeniem, że zachodzi ono w mniejszym stopniu. W przybliżeniu można przyjąć, że kryształ GaAs jest utworzony z jonów $Ga^{+0,2}$ i $As^{-0,2}$. Na rys. 5/10b widzimy, że rozkład gęstości chmury elektronowej w krysztale GaAs wykazuje wyraźną asymetrię.

Polaryzacja wiązania uwidocznia się jeszcze silniej w selenku cynku (ZnSe) – rys. 5/10c. Pierwiastki tworzące ten związek różnią się jeszcze bardziej elektroujemnością, ponieważ cynk należy do II, selen zaś do VI grupy układu okresowego.

Wartości współczynnika jonowości dla kilkunastu związków dwuskładnikowych podano w tabl. 5/1. Symbol oznaczający typ związku stanowi zestawienie numerów grup układu okresowego, do których należą pierwiastki tworzące dany związek. Przykładowo, omawiany arsenek galu jest związkiem typu III–V, a selenek cynku – związkiem typu II-VI.

Z podanego zestawienia wynika, że największe wartości współczynnika jonowości wykazują związki pierwiastków należących do 2 skrajnych grup układu okresowego, tzn. związki typu I–VII. Jednakże nawet w tym przypadku współczynnik jonowości jest mniejszy od jedności. Oznacza to, że nie istnieją kryształy o czysto jonowym wiązaniu. Nawet w takich kryształach, jak rozpatrywana

5/1
Współczynnik jonowości w skali Phillipsa niektórych związków dwuskładnikowych

Typ związku	Symbol chemiczny	Struktura krystaliczna	Współczynnik jonowości
IV-IV	SiC	blendy cynkowej lub wurcytu	0,177
III-V	GaSb	blendy cynkowej	0,261
	GaAs	blendy cynkowej	0,310
	InAs	blendy cynkowej	0,357
	GaP	blendy cynkowej	0,374
	InP	blendy cynkowej	0,421
	GaN	wurcytu	0,500
II-VI	BeO	wurcytu	0,602
	ZnS	blendy cynkowej lub wurcytu	0,623
	CdS	blendy cynkowej lub wurcytu	0,685
	CdO	soli kuchennej	0,785
	MgO	soli kuchennej	0,786
I-VII	CuBr	blendy cynkowej lub wurcytu	0,735
	AgJ	blendy cynkowej lub wurcytu	0,770
	AgCl	soli kuchennej	0,856
	NaCl	soli kuchennej	0,935
	RbF	soli kuchennej	0,960

Źródło: A. J. Van Vechten: Quantum dielectric theory of electronegativity in covalent systems. *Phys. Rev.* **187**, 1007 (1969).

w p. 5.2 sól kuchenna, nie można w pełni wyjaśnić budowy krystalicznej na podstawie sił kulombowskich między jonami Na^+ i Cl^-. Oprócz nich działają siły wynikające z nakrywania się chmur elektronowych tych jonów. Są one skierowane równolegle do krawędzi komórki elementarnej, dzięki czemu tworzy się sieć regularna kryształu NaCl.

Dla kryształów dwuskładnikowych stwierdzono, że związki o współczynniku jonowości większym niż 0,785 krystalizują w strukturze soli kuchennej (lub zbliżonej do niej strukturze chlorku cezu). Oznacza to, że każdy atom ma 6 (lub 8) najbliższych sąsiadów, a wiązania międzyatomowe mają zdecydowanie jonowy charakter. Kryształy te są to typowe dielektryki, nie wykazujące przewodnictwa elektronowego nawet w podwyższonych temperaturach.

Związki o współczynniku jonowości mniejszym niż 0,785 krystalizują natomiast w strukturze blendy cynkowej lub wurcytu[1]. W obu przypadkach każdy atom ma 4 najbliższych sąsiadów, a wiązania międzyatomowe mają zdecydowanie kowalencyjny charakter (kierunkowość wiązań, uwspólnienie par elektronów walencyjnych itd.). Związki te są półprzewodnikami, przy czym większość spośród podanych w tabl. 5/1 jest wykorzystywana w produkowanych obecnie elementach półprzewodnikowych.

5.6 METALE

Wśród wszystkich pierwiastków najliczniejszą grupę stanowią metale. Na rysunku 5/12 przedstawiono układ okresowy pierwiastków z zaznaczonymi pierwiastkami zaliczanymi do metali.

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	J	Xe
Cs	Ba	Lantanowce	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Aktynowce															

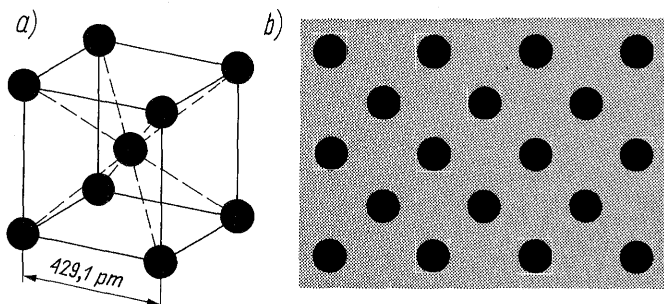
5/12 Metale w układzie okresowym pierwiastków

[1] Wurcyt jest to mało rozpowszechniona odmiana siarczku cynku. Struktura wurcytu należąca do układu heksagonalnego nie będzie tu rozpatrywana.

Teoria wiązań metalicznych jest skomplikowana. Z tego względu zapoznamy się tu jedynie z bardzo uproszczonym modelem metalu, który stosunkowo dobrze pasuje do metali alkalicznych i metali ziem alkalicznych, gorzej natomiast do metali dalszych grup układu okresowego. Model ten rozpatrzymy na przykładzie sodu.

Atom sodu, jak już wyżej zaznaczono, ma 1 elektron walencyjny dość słabo związany z rdzeniem. Gdy atomy Na tworzą kryształ, siły wzajemnego oddziaływania wszystkich rdzeni atomowych i wszystkich elektronów walencyjnych są tak duże, że następuje oderwanie elektronów walencyjnych od ich macierzystych rdzeni. Uwolnione elektrony mogą się prawie swobodnie poruszać w całym krysztale. Stanowią one uwspólniony zbiór, który w postaci rozmytej chmury elektronowej wypełnia w sposób niemal jednorodny całą objętość kryształu. W tym „morzu" ujemnego ładunku elektrycznego są zanurzone jony Na^+, tworzące regularną sieć krystaliczną – rys. 5/13.

Siły wiązania powstają w krysztale wskutek elektrostatycznego oddziaływania między ujemnie naładowaną chmurą elektronową a dodatnimi jonami.



5/13 Sieć krystaliczna sodu: a) komórka elementarna (rozmiary jonów Na^+ zmniejszono 2-krotnie w stosunku do stałej sieci); b) uproszczony (dwuwymiarowy) model sieci (na rysunku tym zachowano właściwe proporcje promieni jonów Na^+ w stosunku do odległości między nimi

Przy pewnej wartości odległości międzyatomowych siły te równoważą siły wzajemnego odpychania między jednoimiennymi jonami. Układ znajduje się wówczas w stanie równowagi, a jego energia przyjmuje wartość minimalną.

Wyrażając się niezbyt ściśle, można powiedzieć, że działanie ściągające chmury elektronowej wiąże w trwały układ odpychające się wzajemnie jony metalu. Z opisu tego wynika, że w przeciwieństwie do wiązania kowalencyjnego, wiązanie metaliczne jest całkowicie zdelokalizowane i bezkierunkowe. Ta bezkierunkowość wiązań sprawia, że jony metalu mogą się dość łatwo przesuwać względem siebie. Za pomocą stosunkowo niewielkich sił zewnętrznych można więc dokonywać plastycznego odkształcenia bryły metalu, nie powodując jej rozkruszenia. Własność tę nazywamy *plastycznością* (lub kowalnością) metalu.

Sód krystalizuje w układzie regularnym przestrzennie centrowanym (rys. 5/13a). Długość wiązania wynosi 372 pm, a promień jonu Na^+ – 98 pm.

Wynika stąd, że jony sodu zajmują tylko nieznaczną część (ok. 10%) objętości kryształu.

W sodzie i innych metalach alkalicznych za wiązania międzyatomowe odpowiedzialne są prawie wyłącznie oddziaływania między rdzeniami atomowymi, a elektronami walencyjnymi. Odległości międzyatomowe są wówczas dość duże, a zatem gęstość tych metali jest niewielka, a wiązania są stosunkowo słabe[1].

W metalach przejściowych występują dodatkowe siły wiązania, wynikające z oddziaływań między wewnętrznymi powłokami elektronowymi. Jony tych metali mają większe rozmiary. W krysztale można wyobrażać je sobie jako stykające się ze sobą kule. Metale te charakteryzują się dużymi energiami wiązania i w związku z tym wysokimi temperaturami topnienia.

Koncentracja swobodnych elektronów w metalu jest bardzo duża. Dla większości metali zawiera się ona w granicach 1–2 elektronów na atom, a więc jest rzędu 10^{23} cm^{-3}. Dzięki dużej koncentracji swobodnych elektronów metale odznaczają się dobrym przewodnictwem elektrycznym i cieplnym. Swobodne elektrony w metalu nazywamy w związku z tym *elektronami przewodnictwa*.

Swoboda ruchu elektronów wewnątrz metalu upodobnia je w pewnym stopniu do cząsteczek gazu w zamkniętym naczyniu. Z tego względu zbiór wszystkich swobodnych elektronów w metalu nazywamy często *gazem elektronowym*.

---

[1] Energia wiązania metali alkalicznych jest znacznie mniejsza niż halogenków tych metali.

6

STRUKTURA ENERGETYCZNA CIAŁA STAŁEGO

W poprzednim rozdziale stwierdziliśmy, że ciała stałe, ze względu na ich przewodnictwo elektronowe, mogą być podzielone na metale, półprzewodniki i dielektryki. Podział na te 3 klasy, jak widzieliśmy, jest uwarunkowany rodzajem wiązań międzyatomowych w krysztale. Klasyfikacji tej nie można jednak dokonać w sposób precyzyjny, albowiem granice między poszczególnymi klasami nie są ściśle określone. Najmniej wątpliwości budzi wyodrębnienie metali, natomiast podział pozostałych ciał stałych na półprzewodniki i dielektryki jest w znacznym stopniu sprawą umowną.

W rozdziale tym zapoznamy się z tzw. *teorią pasmową* ciał stałych. Jest to dział fizyki, zajmujący się badaniem struktury poziomów energetycznych elektronów w kryształach. Teoria ta stanowi podstawę rozwoju współczesnej fizyki metali, półprzewodników i dielektryków, umożliwia bowiem wyjaśnienie najważniejszych własności elektrycznych i optycznych tych materiałów.

6.1 ROZSZCZEPIENIE POZIOMÓW ENERGETYCZNYCH W CIELE STAŁYM

6.1.1 Metale

Przy omawianiu wiązania występującego w układzie złożonym z 2 atomów wodoru widzieliśmy, że stan podstawowy układu składa się z 2 poziomów energetycznych, które powstają w wyniku rozszczepienia poziomu $1s$ pojedynczego atomu wodoru. Każdy z tych poziomów może być obsadzony przez 2 elektrony. Z faktu tego można wyciągnąć wniosek, że dla układu złożonego z N atomów wodoru stan podstawowy zawiera N poziomów energetycznych, z których każdy może pomieścić 2 elektrony. W układzie tym występuje więc łącznie $2N$ stanów kwantowych, powstałych wskutek rozszczepienia poziomu $1s$.

Ogólnie, zbliżanie do siebie N jednakowych atomów na odległość, przy której funkcje falowe ich elektronów zaczynają zachodzić na siebie, prowadzi, ze względu na zakaz Pauliego, do rozszczepienia każdego z atomowych poziomów energetycznych na N blisko siebie położonych dyskretnych poziomów.

Układem złożonym z N jednakowych blisko siebie położonych atomów jest kryształ danego pierwiastka. Nie będziemy tu rozpatrywali pojęciowo najprostszego kryształu wodoru atomowego, ponieważ pod normalnym ciśnieniem wodór tworzy kryształy molekularne, zbudowane z cząsteczek H_2 i dopiero pod ciśnieniem rzędu milionów atmosfer jest spodziewane powstanie wodoru metalicznego. Tego jednakże nie udało się jeszcze zrealizować. Rozważymy natomiast strukturę energetyczną dobrze znanego metalu, jakim jest sód.

Atom sodu ma konfigurację elektronową $1s^2 2s^2 2p^6 3s^1$, zawiera więc 4 podpowłoki elektronowe. Dla położonego najbliżej jądra elektronu $1s$ można oszacować najbardziej prawdopodobną odległość od jądra, korzystając ze wzoru (3/5). Należy jedynie uwzględnić to, że w atomie wodoru, dla którego został wyprowadzony ten wzór, ładunek jądra wynosi q, tu zaś jest on równy Zq. Siła, z jaką jądro atomu sodu przyciąga elektron $1s$ jest więc Z razy większa, a odpowiedni promień Z razy mniejszy niż w atomie wodoru. Podstawiając dla sodu $Z = 11$ i dla elektronu $1s$ $n = 1$, znajdziemy

$$r_{1s} = \frac{a_0}{11} = 4{,}8 \text{ pm} \tag{6/1}$$

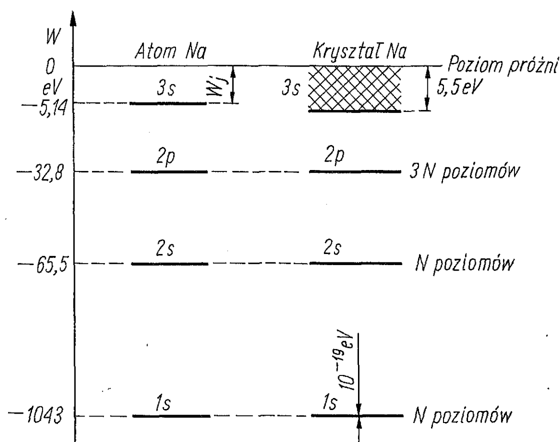
Z kolei najbardziej zewnętrzny elektron $3s$ w atomie sodu porusza się w polu wytwarzanym przez jądro (zawierające 11 protonów) i 10 elektronów powłok wewnętrznych. Wypadkowy ładunek elektryczny, oddziałujący na elektron $3s$, jest więc równy $+q$, tak jak to występuje w atomie wodoru. Przy wyznaczaniu najbardziej prawdopodobnej odległości tego elektronu od jądra, możemy więc posługiwać się modelem *atomu wodoropodobnego*. Podstawiając we wzorze (3/5) $n = 3$, otrzymamy dla elektronu $3s$

$$r_{3s} = 9a_0 = 480\, pm \tag{6/2}$$

Odległość między sąsiednimi atomami w krysztale sodu, jak wiemy, wynosi 372 pm. Z porównania tych liczb wynika, że funkcje falowe elektronów $1s$ prawie nie zachodzą na siebie, a więc rozszczepienie poziomu $1s$ w krysztale sodu jest znikomo małe. Dokładne obliczenia prowadzą do wniosku, że rozszczepienie to jest rzędu 10^{-19} eV.

Podobnie przedstawia się sprawa dla elektronów $2s$ i $2p$ z tym, że promienie r_{2s} i r_{2p} są odpowiednio większe od r_{1s}, ale wciąż znacznie mniejsze niż połowa odległości między sąsiednimi atomami w krysztale. W wyniku tego rozszczepienie poziomów $2s$ i $2p$ jest również pomijalnie małe.

Istotną różnicę obserwujemy natomiast w przypadku elektronów $3s$. Promień r_{3s} jest tak duży, że funkcje falowe tych elektronów zachodzą na siebie w znacznym stopniu, a w wyniku tego rozszczepienie poziomu $3s$ jest bardzo duże – zob. rys. 6/1.



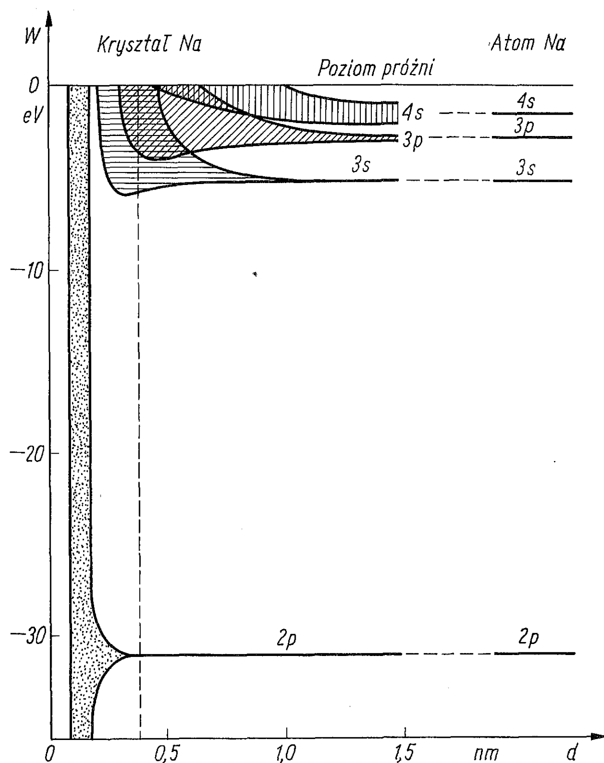
6/1 Poziomy energetyczne w atomie sodu i w krysztale sodu (wykres poglądowy, bez zachowania skali na osi energii)
W_j — energia jonizacji atomu sodu

Ponieważ w 1 cm³ kryształu sodu znajduje się $2{,}65 \cdot 10^{22}$ atomów tego pierwiastka, zatem rozszczepione poziomy są w skali energii rozmieszczone niezwykle gęsto. W większości rozważań przyjmuje się, że tworzą one *pasmo energetyczne*, w którym energia może zmieniać się w sposób ciągły. Górnym ograniczeniem tego pasma jest tzw. *poziom próżni*, któremu przypisujemy energię równą zeru. Elektron o energii większej od tej wartości jest uwolniony z sił wiążących go z atomem lub kryształem i może swobodnie poruszać się w otaczającej przestrzeni (tzn. w próżni).

Proces rozszczepiania się atomowych poziomów energetycznych w pasma przedstawiono na rys. 6/2. Pokazany tam wykres dotyczy hipotetycznego eksperymentu, polegającego na ściskaniu i rozciąganiu kryształu w taki sposób, że wszystkie odległości *d* między sąsiednimi atomami ulegają zmianom w tym samym stosunku. Jeżeli odległości te są dostatecznie duże, to kryształ zachowuje się tak, jak zbiór nie oddziałujących ze sobą atomów, a więc rozszczepienie poziomów energetycznych praktycznie nie występuje. W miarę zmniejszania odległości *d* jest widoczne coraz silniejsze oddziaływanie zachodzące między atomami kryształu.

Na rysunku 6/2 uwzględniono nie tylko poziom 3*s*, lecz również wyższe poziomy energetyczne, które w stanie podstawowym atomu sodu nie są obsadzone. Poziomy te wnoszą swój wkład do pasma energetycznego, które powstaje z rozszczepienia poziomu 3*s*. Aby niepotrzebnie nie komplikować rozważań, w dalszych wywodach nie będziemy tych dodatkowych stanów kwantowych brali pod uwagę. W przypadku sodu i wielu innych metali nie powoduje to błędu w prowadzonym rozumowaniu.

Rozszczepienie poziomów atomowych w pasma energetyczne ma bardzo ważne konsekwencje w postaci określonych własności fizycznych (zwłaszcza elektrycznych i optycznych) kryształu. Zanim jednak przystąpimy do zapoznania się



6/2 Rozszczepianie się atomowych poziomów energetycznych przy tworzeniu się kryształu sodu. Linią przerywaną pokazano rzeczywistą odległość sąsiednich atomów w krysztale

z tymi własnościami, warto zastanowić się nad istotną różnicą w zachowaniu się elektronów należących do poszczególnych pasm energetycznych.

Elektron w atomie może przebywać w danym stanie stacjonarnym nieskończenie długo. W krysztale natomiast czas przebywania elektronu w pobliżu określonego atomu, na mocy zasady nieoznaczoności Heisenberga, jest związany z szerokością ΔW danego pasma energetycznego. Wartość tego czasu można oszacować na podstawie zależności (1/39). Podstawiając dla elektronu 1*s* wartość $\Delta W = 10^{-19}$ eV, dochodzimy do wniosku, że czas przebywania tego elektronu w pobliżu określonego atomu jest rzędu godziny. Elektrony wewnętrznych powłok atomowych można więc uważać za trwale związane z „macierzystym" jądrem atomowym. Są one zatem silnie zlokalizowane i tworzą wraz z jądrami rdzenie atomowe, z których jest zbudowana sieć krystaliczna ciała stałego.

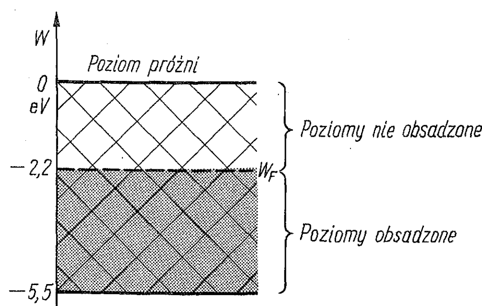
Zupełnie inaczej zachowują się elektrony walencyjne. Ponieważ szerokość pasma powstałego w wyniku rozszczepienia poziomu 3*s* wynosi dla sodu 5,5 eV, zatem czas przebywania tego elektronu w pobliżu określonego atomu jest rzędu 10^{-16} s. Oznacza to, że elektrony walencyjne są całkowicie uwspólnione przez wszystkie atomy kryształu. Tworzą one, jak to już powiedziano w p. 5.6, rozmytą chmurę elektronową, zwaną *gazem elektronowym*, która wiąże odpychające się wzajemnie jony metalu w trwały układ.

Dzięki możliwości swobodnego przemieszczania się w krysztale, elektrony te decydują o takich jego własnościach, jak duże przewodności elektryczna i cieplna oraz duża wartość współczynnika odbicia promieniowania elektromagnetycznego w zakresie widzialnym i w podczerwieni. Z tych względów zajmiemy się teraz dokładniej pasmem 3*s* w rozważanym krysztale sodu.

Sód jest pierwiastkiem jednowartościowym. Każdy atom ma więc 1 elektron walencyjny, a zatem pasmo 3*s* jest obsadzone przez N elektronów. Liczba stanów kwantowych tego pasma wynosi natomiast $2N$, a więc tylko połowa z nich jest obsadzona[1)]. Fakt ten ma bardzo doniosłe konsekwencje. Łatwo można bowiem uzmysłowić sobie, że pasmo całkowicie obsadzone nie daje żadnego przyczynku do przepływu prądu elektrycznego. Aby bowiem w ciele stałym popłynął prąd, musi nastąpić przyspieszenie elektronów pod działaniem sił pola elektrycznego. Przyspieszeniu temu towarzyszy wzrost energii elektronów, tzn. przeniesienie ich z niższych na wyższe poziomy energetyczne. W tym celu, musimy w danym pasmie dysponować nie tylko elektronami, lecz również nie obsadzonymi poziomami energetycznymi. W pasmie całkowicie zapełnionym brak jest takich nie obsadzonych poziomów i z tego względu elektrony tego pasma nie biorą udziału w przepływie prądu elektrycznego przez kryształ.

W krysztale sodu niższe pasma energetyczne (tzn. bardzo wąskie pasma 1*s*, 2*s* i 2*p*) są całkowicie obsadzone przez elektrony i z uwagi na przewodnictwo elektryczne nie interesują nas zupełnie. Pasmo 3*s* natomiast, jak to już zaznaczono, jest obsadzone tylko częściowo.

W temperaturze zera bezwzględnego każdy układ fizyczny znajduje się w stanie o możliwie najmniejszej energii. Elektrony w pasmie 3*s* obsadzają więc w tych warunkach wszystkie najniższe poziomy energetyczne aż do pewnego poziomu granicznego, który nazywamy *poziomem Fermiego* i któremu przypisujemy energię W_F, nazywaną *energią Fermiego*. Obsadzenie pasma 3*s* w krysztale sodu w tej temperaturze pokazano na rys. 6/3.



6/3 Obsadzenie najwyższego pasma energetycznego w krysztale sodu w temperaturze zera bezwzględnego
W_F — energia Fermiego

[1)] W rzeczywistości obsadzenie jest znacznie mniejsze ponieważ w pasmie tym występują dodatkowe, pominięte przez nas stany kwantowe, powstałe z rozszczepienia wyższych poziomów energetycznych atomów sodu.

W temperaturach różnych od zera bezwzględnego podział poziomów energetycznych na obsadzone i nieobsadzone nie przebiega tak ostro. Istnieje wówczas pewien zakres energii, w którym prawdopodobieństwo obsadzenia poszczególnych poziomów zmienia się monotonicznie od wartości praktycznie równej jedności (poziomy obsadzone) do wartości bliskiej zera (poziomy nieobsadzone)[1]. Szerokość tego zakresu jest orientacyjnie równa 8kT. W temperaturze pokojowej ($T = 300$ K) wartość ta wynosi ok. 0,2 eV, jest zatem pomijalnie mała w stosunku do szerokości pasma 3*s*. Można więc w przybliżeniu przyjąć, że wykres przedstawiony na rys. 6/3 odnosi się również do temperatury pokojowej.

Pod działaniem pola elektrycznego elektrony o energii bliskiej W_F mogą zwiększać swą energię, zajmując nie obsadzone wyższe poziomy energetyczne. Mogą one więc brać udział w przewodzeniu prądu elektrycznego. Ze względu na to, że elektronów tych jest bardzo dużo, sód charakteryzuje się dużą przewodnością elektryczną.

To co tutaj powiedziano o sodzie odnosi się z pewnymi mało istotnymi różnicami także do innych metali. Można więc powiedzieć, że *cechą charakterystyczną metali, odróżniającą je np. od półprzewodników, jest to, że ich struktura energetyczna zawiera pasmo, które w temperaturze zera bezwzględnego jest tylko częściowo obsadzone przez elektrony*. W wyniku tego, metale są dobrymi przewodnikami prądu elektrycznego. Wykazują one przy tym przewodnictwo w temperaturze zera bezwzględnego. Jak zobaczymy w p. 10.5, przewodność metali maleje ze wzrostem temperatury, a zatem w temperaturze zera bezwzględnego osiąga ona swą największą wartość.

6.1.2 Półprzewodniki

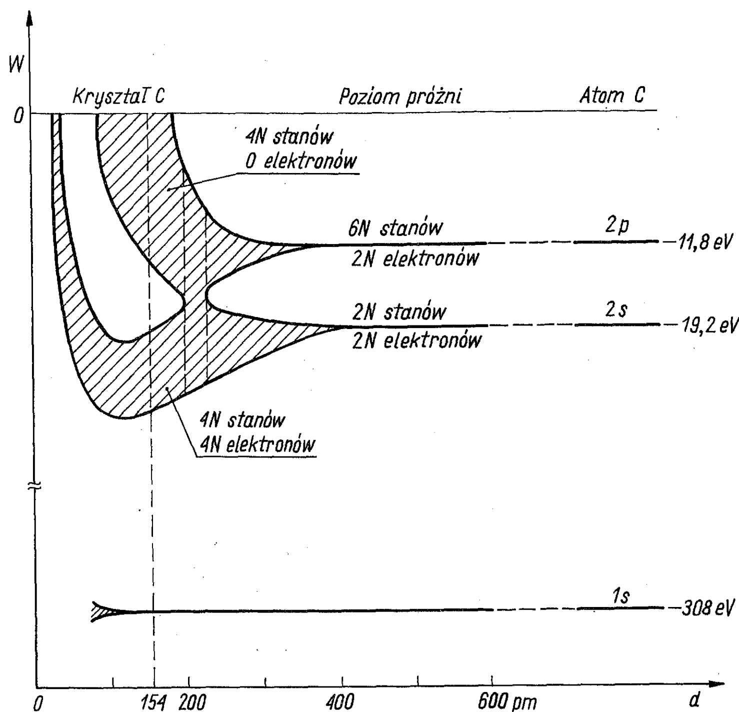
Przejdźmy teraz do bardziej złożonego przypadku, jakim są struktury energetyczne pierwszych 3 pierwiastków IV grupy układu okresowego, tzn. węgla, krzemu i germanu. Pierwiastki te krystalizują w strukturze diamentu i mają podobne własności. Dla ustalenia uwagi, zajmiemy się początkowo najlżejszym z tych pierwiastków, tzn. węglem, o konfiguracji elektronowej $1s^2 2s^2 2p^2$. Atom węgla ma więc całkowicie zapełnioną powłokę *K*, jego powłoka *L* jest natomiast tylko częściowo obsadzona przez 2 elektrony 2*s* i 2 elektrony 2*p*.

Przy tworzeniu się kryształu diamentu poziomy atomowe ulegają rozszczepieniu w pasma energetyczne, przy czym rozszczepienie poziomu 1*s* (analogicznie jak w przypadku sodu) jest znikomo małe. Przeciwnie zaś, poziomy 2*s* i 2*p*, położone stosunkowo blisko siebie, ulegają znacznemu rozszczepieniu, podobnie jak to było z poziomem 3*s* w krysztale sodu.

Proces rozszczepiania się poziomów atomowych węgla w pasma przedstawiono na rys. 6/4. Rysunek ten jest odpowiednikiem rys. 6/2 i podobnie jak tamten odnosi się do hipotetycznego eksperymentu z izotropowo ściskanym i rozcią-

---

[1] Prawdopodobieństwem obsadzenia danego poziomu przez elektron zajmuje się *statystyka Fermiego-Diraca*. Zagadnienia te będą bardziej szczegółowo omawiane w p. 8.2.



6/4 Rozszczepianie się atomowych poziomów energetycznych przy tworzeniu się kryształu diamentu (wykres poglądowy, bez zachowania skali na osi energii)

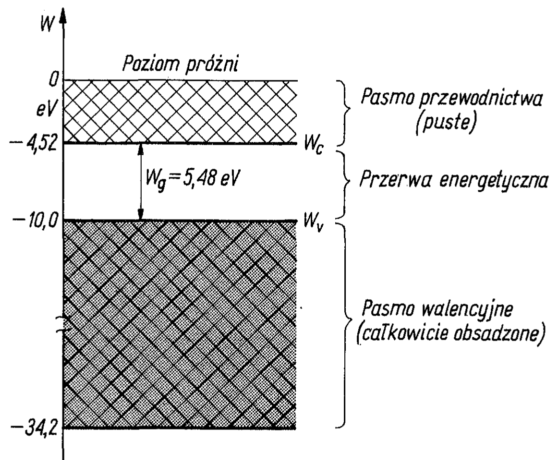
ganym kryształem. Na rysunku 6/4 można zaobserwować, że w miarę zmniejszania odległości międzyatomowych pasma 2*s* i 2*p* początkowo rozszczepiają się, następnie łączą w jedno pasmo i wreszcie rozdzielają na 2 niezależne pasma, przy czym każde z nich zawiera taką samą liczbę stanów kwantowych. Pasmo położone niżej zawiera 4*N* stanów wiążących, wyższe zaś – 4*N* stanów antywiążących[1]. Między tymi pasmami pojawia się pewien obszar energii zabronionych dla elektronu. Nosi on nazwę *przerwy energetycznej*. Szerokość tej przerwy oznaczamy symbolem W_g.

Atom węgla ma 4 elektrony walencyjne. W temperaturze zera bezwzględnego obsadzają one najniżej położone poziomy energetyczne. Obsadzają więc całkowicie niższe pasmo energetyczne, które w związku z tym nazywamy *pasmem walencyjnym*. Wyższe, całkowicie nieobsadzone pasmo nosi nazwę *pasma przewodnictwa*.

Pasma te oraz stopień ich obsadzenia w temperaturze zera bezwzględnego pokazano na rys. 6/5. Najwyższy poziom energetyczny w pasmie walencyjnym nazywamy *wierzchołkiem* tego pasma, a odpowiadającą mu energię oznaczamy

---

[1] Mieszanie się stanów kwantowych z pasm 2*s* i 2*p* oraz wytwarzanie się stanów wiążących i antywiążących są w ścisłym związku z powstawaniem zhybrydyzowanego orbitalu sp^3, w którym była mowa w p. 5.4 przy omawianiu wiązań atomowych w strukturze diamentu.



6/5 Schemat pasm energetycznych diamentu w temperaturze zera bezwzględnego

symbolem W_v. Podobnie, najniższy poziom w pasmie przewodnictwa jest nazywany *dnem* tego pasma, a odpowiadająca mu energia – oznaczana symbolem W_c. Zgodnie z definicją, szerokość przerwy energetycznej wynosi więc[1)]

$$\boxed{W_g = W_c - W_v} \qquad (6/3)$$

Ponieważ ani pasmo całkowicie obsadzone, ani pasmo puste nie biorą udziału w przewodzeniu prądu elektrycznego, zatem diament w temperaturze zera bezwzględnego jest izolatorem.

W wyższych temperaturach następuje pod wpływem energii cieplnej wzbudzenie elektronów na wyższe poziomy energetyczne. Pewna ich liczba zostaje więc przeniesiona z pasma walencyjnego do pasma przewodnictwa. W pasmie przewodnictwa pojawiają się wskutek tego elektrony, które mogą w krysztale wytwarzać prąd elektryczny. Jednocześnie w pasmie walencyjnym powstają nieobsadzone stany kwantowe, co także umożliwia istniejącym tam elektronom udział w przepływie prądu. Im wyższa jest temperatura, tym większa liczba elektronów zostaje wzbudzona z pasma walencyjnego do pasma przewodnictwa, tym większa jest zatem przewodność kryształu.

Innym sposobem wzbudzenia elektronów w diamencie (i innych kryształach o podobnej strukturze energetycznej) jest poddanie go działaniu promieniowania elektromagnetycznego o dostatecznie dużej energii fotonów. Jeżeli energia fotonu jest większa niż szerokość przerwy energetycznej, tzn. jeżeli jest spełniony warunek

$$h\nu > W_g$$

to może nastąpić absorpcja tego fotonu w krysztale, przy czym zaabsorbowana energia zostaje zużyta na przeniesienie 1 elektronu z pasma walencyjnego do pasma przewodnictwa. Kryształ może wówczas przewodzić prąd elektryczny na-

[1)] Użyte tu indeksy *v*, *c* i *g*, powszechnie stosowane w literaturze, pochodzą od terminów angielskich: *valence band* (pasmo walencyjne), *conduction band* (pasmo przewodnictwa) i *energy gap* (przerwa energetyczna).

wet w temperaturze zera bezwzględnego. Zjawisko pojawienia się przewodnictwa elektrycznego (a ogólniej – zmiany przewodności elektrycznej) wskutek absorpcji promieniowania elek magnetycznego nazywamy *fotoprzewodnictwem*.

Jak już powiedziano w p. 5.4, kryształy wykazujące omawiane tutaj własności nazywamy *półprzewodnikami*. *Cechą charakterystyczną półprzewodników jest to, że ich struktura energetyczna w temperaturze zera bezwzględnego składa się z pasm całkowicie obsadzonych oraz pasma pustego*, tj. takiego, w którym żaden stan kwantowy nie jest obsadzony. W wyniku tego półprzewodniki nie poddane działaniu promieniowania elektromagnetycznego nie wykazują w temperaturze zera bezwzględnego przewodnictwa elektrycznego (są izolatorami). Stają się one natomiast przewodnikami prądu elektrycznego pod wpływem nagrzania do dostatecznie wysokiej temperatury lub absorpcji promieniowania o dostatecznie dużej energii fotonów. W przeciwieństwie do metali przewodność czystych (niedomieszkowanych) półprzewodników wzrasta ze wzrostem temperatury.

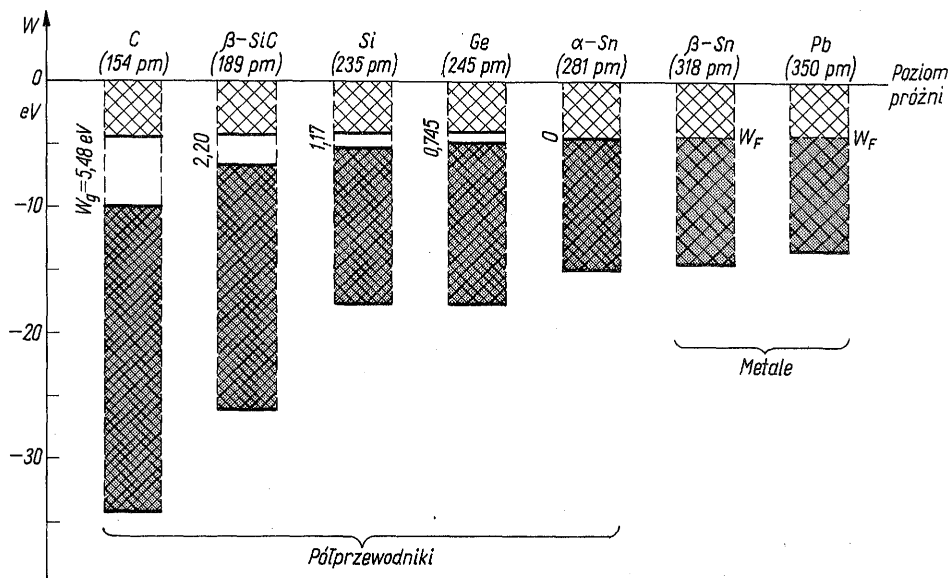
Podana wyżej definicja może budzić pewne zastrzeżenia, ponieważ zgodnie z nią, diament, traktowany przez większość autorów jako dielektryk, został zaliczony do klasy półprzewodników. Ma on jednakże wszelkie cechy półprzewodnika, a różni się od krzemu i germanu znacznie większą szerokością przerwy energetycznej. W diamencie występuje zjawisko fotoprzewodnictwa, które nawet praktycznie stosowano do detekcji fotonów o dużej energii[1].

W literaturze często za podstawę podziału substancji niemetalicznych na półprzewodniki i dielektryki przyjmuje się wartość przerwy energetycznej. Ciała o $W_g < 2.5$ eV zalicza się do półprzewodników, natomiast o większej szerokości przerwy energetycznej – do dielektryków. Ta dość rozpowszechniona definicja nie bierze pod uwagę odmienności zjawisk fizycznych występujących w tych 2 klasach materiałów, a opiera się jedynie na wartości liczbowej parametru, od którego m.in. zależy przewodność materiału w określonej temperaturze.

Definicja ta jest ponadto chybiona. Istnieją bowiem materiały o przerwie energetycznej większej niż 2,5 eV (np. ZnS – 3,56 eV, GaN – 3,44 eV), które mają wszelkie typowe cechy półprzewodników i które znajdują zastosowanie w produkowanych elementach półprzewodnikowych. Ponadto dla wielu dielektryków (np. siarki, jodu i innych kryształów molekularnych) pojęcie przerwy energetycznej może w ogóle nie mieć sensu.

Dla pozostałych pierwiastków IV grupy układu okresowego, krystalizujących w strukturze diamentu, wykres przedstawiający mieszanie się 2 najwyższych pasm energetycznych ma charakter podobny do podanego na rys. 6/4. Występują tu jedynie różnice ilościowe wywołane różnymi wartościami promieni atomowych i odległości międzyatomowych w poszczególnych kryształach. Odległości między sąsiednimi atomami wzrastają w miarę przechodzenia ku pierwiastkom o coraz większej liczbie atomowej. W wyniku tego zmienia się szerokość przerwy energetycznej – zob. rys. 6/6.

[1] Zob. np.: L. F. Curtiss, B. W. Brown: Diamond as a gamma-ray counter. *Phys. Rev.* **72**, 643 (1947).



6/6 Schematy pasm energetycznych kryształów pierwiastków należących do IV grupy układu okresowego w temperaturze zera bezwzględnego. Pod symbolami podano odległość między sąsiednimi atomami w krysztale

Na rysunku tym uwzględniono nie tylko węgiel, krzem i german, lecz również węglik krzemu (β-SiC) – półprzewodnik krystalizujący w strukturze blendy cynkowej, oraz pozostałe pierwiastki należące do IV grupy układu okresowego, tj. cynę i ołów.

Cyna, jak już wspomniano w p. 5.4, ma 2 odmiany alotropowe. Cyna szara (α-Sn) krystalizuje w strukturze diamentu i jest półprzewodnikiem o dość specyficznych własnościach (jest to półprzewodnik o zerowej przerwie energetycznej – zob. p. 7.5), natomiast cyna biała (β-Sn) krystalizuje w układzie tetragonalnym i podobnie jak ołów (układ regularny płasko centrowany) jest metalem. W metalach pasmo walencyjne zlewa się z pasmem przewodnictwa w jedno wspólne pasmo, które w temperaturze zera bezwzględnego jest zapełnione aż do poziomu W_F.

Dla typowych dielektryków, jakimi są np. kryształy jonowe, schemat pasm energetycznych i ich obsadzenie przez elektrony przedstawiają się w sposób analogiczny do tego, jaki występuje w półprzewodnikach, tzn. taki, jak na rys. 6/5. Przerwy energetyczne mają w tych kryształach duże wartości, zawierające się zwykle w granicach 5–10 eV. Kryształów tych nie zaliczamy jednak do półprzewodników nie ze względu na wartość przerwy energetycznej, ale dlatego, że nie wykazują one przewodnictwa *elektronowego.* W podwyższonych temperaturach może się w nich pojawić przewodnictwo jonowe, a po przekroczeniu pewnej temperatury kryształy te ulegają stopieniu.

Fakt istnienia przerwy energetycznej w półprzewodnikach i w większości dielektryków odbija się w bardzo charakterystyczny sposób na ich własnościach optycznych. Dla fotonów o energii mniejszej niż W_g czysty (niedomieszkowany) półprzewodnik jest całkowicie przezroczysty, natomiast fotony o energii większej niż W_g są silnie pochłaniane, w związku z czym kryształ staje się nieprzezroczysty. Jeżeli granica pochłaniania fotonów, nazywana *krawędzią absorpcji*, leży w za-

kresie widzialnym, to kryształ przepuszcza jedynie część widma świetlnego, wskutek czego ma on charakterystyczne zabarwienie, zależne od szerokości przerwy energetycznej. Człowiek dysponuje więc narządem zmysłowym, za pomocą którego może w pewnym zakresie ocenić szerokość przerwy energetycznej czystych półprzewodników.

Problem 6.1. Wyznacz warunek na długość fali promieniowania elektromagnetycznego pochłanianego przez czysty półprzewodnik o przerwie energetycznej W_g [eV]. Stwierdź, które z półprzewodników zestawionych w dodatku D są przezroczyste, nieprzezroczyste lub częściowo przezroczyste dla światła (zakres widzialny obejmuje promieniowanie o długości od 380 do 760 nm). Zastanów się, jak zmieniałaby się barwa światła przepuszczanego przez kryształ, gdyby jego przerwę energetyczną stopniowo zwiększać od 1 do 5 eV.

6.2 ELEKTRON W POLU PERIODYCZNYM

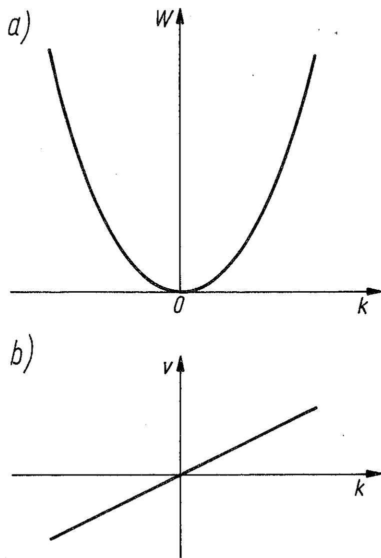
Rozważania zawarte w poprzednim punkcie ograniczały się do zagadnień związanych z wartościami energii, jakie może mieć elektron w krysztale. Stwierdziliśmy tam, że energie te tworzą pewne quasi-ciągłe pasma, przedzielone zakresami energii zabronionych – przerwami energetycznymi. Dyskutowaliśmy również sprawę obsadzenia tych pasm w metalach i w półprzewodnikach. Na tej podstawie mogliśmy wyjaśnić przyczyny odmiennego zachowania się elektronów poddanych działaniu pola elektrycznego w tych 2 klasach materiałów. Istnieje jednak szereg ważnych własności elektrycznych i optycznych, których nie można wytłumaczyć na podstawie rozpatrywanych w p. 6.1 schematów pasm energetycznych. Nie można na przykład odpowiedzieć na pytanie, dlaczego z arsenku galu można zbudować laser półprzewodnikowy, a z krzemu nie można tego zrobić, pomimo tego, że oba te półprzewodniki mają podobną strukturę krystaliczną i zbliżony schemat pasm energetycznych.

Opis zachowania się elektronów w półprzewodniku za pomocą schematu pasm energetycznych jest więc w wielu przypadkach niewystarczający. Wynika to stąd, że do pełnego scharakteryzowania ruchu elektronu w ciele stałym jest konieczna znajomość nie tylko jego energii, lecz również jego wektora falowego. Energia elektronu w ciele stałym jest w dość skomplikowany sposób uzależniona od wektora falowego. Dalsze punkty tego rozdziału będą poświęcone poznaniu tej zależności i wypływających z niej wniosków dla jednowymiarowej struktury periodycznej, będącej bardzo grubym przybliżeniem rzeczywistego kryształu.

Zanim jednak przystąpimy do analizy ruchu elektronu w polu periodycznym, poświęcimy nieco uwagi zależności $W(k)$ dla elektronu swobodnego (w próżni). Zależność tę, wyrażoną wzorem (2/13) przedstawiono na rys. 6/7a. Jest to relacja nieliniowa i z tego względu jest ona często nazywana *zależnością dyspersyjną*[1].

---

[1] *Dyspersja fal* jest to zależność prędkości fazowej fal harmonicznych od ich częstotliwości. Dla fal nie wykazujących dyspersji (np. dla fal elektromagnetycznych w próżni) długość fali jest odwrotnie proporcjonalna do częstotliwości. Dla rozpatrywanych tu fal de Broglie'a nie zachodzi natomiast proporcjonalność między $W = h\nu$, a $k = 2\pi/\lambda$.



6/7
Zależność dyspersyjna (a) i zależność prędkości od wektora falowego (b) dla elektronu swobodnego (w próżni)

Różniczkując funkcję $W(k)$, otrzymamy

$$\frac{\mathrm{d}W}{\mathrm{d}k} = \frac{\hbar^2 k}{m} = \hbar v \tag{6/4}$$

gdzie v – prędkość elektronu. Ostatnia równość wynika z relacji de Broglie'a, zgodnie z którą iloczyn $\hbar k$ jest równy pędowi mv elektronu. Na podstawie wzoru (6/4) można napisać

$$v = \frac{1}{\hbar} \frac{\mathrm{d}W}{\mathrm{d}k} \tag{6/5}$$

Otrzymana zależność, jak łatwo można sprawdzić, pokrywa się z definicją prędkości grupowej – zob. wzór (1/23).

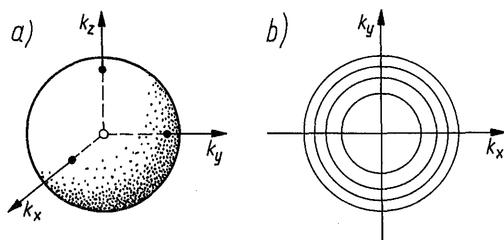
Z wykresu $W(k)$ można więc odczytać prędkość, z jaką porusza się elektron. Zależność tej prędkości od wektora falowego pokazano na rys. 6/7b.

Następne różniczkowanie funkcji $W(k)$ prowadzi do wniosku, że

$$\frac{1}{m} = \frac{1}{\hbar^2} \frac{\mathrm{d}^2 W}{\mathrm{d}k^2} \tag{6/6}$$

Wykres przedstawiony na rys. 6/7a zawiera więc również informację dotyczącą masy poruszającej się cząstki. Im smuklejsza jest parabola reprezentująca zależność $W(k)$, tym mniejsza jest masa rozpatrywanej cząstki. W rozpatrywanym przypadku masa elektronu jest stała, niezależna od jego energii, ponieważ, zgodnie z przyjętym założeniem, nie uwzględniamy tu efektów relatywistycznych.

Ze wzoru (2/13) wynika jeszcze jeden ważny wniosek. Na jego podstawie można mianowicie określić kształt tzw. *powierzchni izoenergetycznej*, tj. powierzchni stałej wartości energii w przestrzeni, w której współrzędnymi są składowe k_x, k_y i k_z wektora falowego. Dla elektronu swobodnego jest nią sfera, czyli powierzchnia kuli, przedstawiona na rys. 6/8. Wynika to z faktu, że dla elektro-



6/8 Powierzchnia izoenergetyczna dla elektronu swobodnego: a) rzut perspektywiczny; b) linie $W = \text{const}$ w płaszczyźnie $k_x k_y$

nu swobodnego energia nie zależy od kierunku wektora falowego, a jedynie od jego wartości.

Przejdźmy teraz do analizy zachowania się elektronu w krysztale. Jest to zagadnienie tak skomplikowane, że jego rozwiązanie bez daleko idących uproszczeń jest niemożliwe.

Jako pierwsze uproszczenie przyjmiemy, że rdzenie atomowe, stanowiące osnowę kryształu, są nieruchome i tworzą idealnie uporządkowaną sieć przestrzenną[1]. W polu wytwarzanym przez te rdzenie porusza się olbrzymia ilość wzajemnie ze sobą oddziałujących elektronów walencyjnych. Problem ruchu tych elektronów jest więc „nieskończenie" bardziej złożony niż w atomie wieloelektronowym. W celu rozwiązania tego problemu, zastosujemy tu, podobnie jak to czyniliśmy w p. 3.4, *przybliżenie jednoelektronowe*. Polega ono na tym, że zamiast rozpatrywać oddziaływanie danego elektronu z każdym z pozostałych elektronów walencyjnych, przyjmiemy, że rozpatrywany elektron porusza się w pewnym polu wypadkowym, wytwarzanym przez wszystkie jądra atomowe i wszystkie pozostałe elektrony. Założymy przy tym, że potencjał tego pola, zwany *potencjałem efektywnym* jest niezależny od czasu.

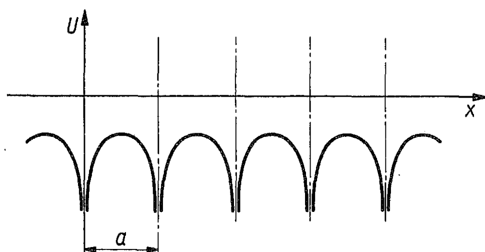
W idealnym nieskończonym krysztale potencjał ten ma pełną symetrię sieci krystalicznej, w szczególności wykazuje on określoną periodyczność wzdłuż każdej prostej sieciowej.

Mimo znacznego uproszczenia, jakie uzyskuje się dzięki zastosowaniu przybliżenia jednoelektronowego, rozwiązanie równania Schrödingera dla obranego elektronu pozostaje nadal sprawą bardzo skomplikowaną, m.in. ze względu na to, że rozkład potencjału efektywnego nie jest z góry określoną funkcją.

W celu dalszego uproszczenia zagadnienia, ograniczymy nasze rozważania do przypadku jednowymiarowego. Potencjał efektywny ma wówczas przebieg, który w sposób poglądowy przedstawiono na rys. 6/9[2]. Linie przerywane reprezentują tu płaszczyzny sieciowe, w których potencjał staje się nieskończenie wielki. Płaszczyzny te dzielą całą przestrzeń na nieskończony ciąg identycznych

[1] W rzeczywistości wskutek posiadanej energii cieplnej rdzenie te wykonują drgania wokół swych położeń równowagi.

[2] Na osi rzędnych odłożono tu energię potencjalną elektronu, która, jak wynika ze wzoru (1/8), jest proporcjonalna do potencjału ze znakiem minus.



6/9 Rozkład energii potencjalnej w jednowymiarowym modelu kryształu

komórek, rozciągających się w nieskończoność w kierunkach prostopadłych do osi x.

Zgodnie z zastosowanym przybliżeniem, w polu przedstawionym na rys. 6/9 porusza się 1 elektron. Znając funkcję $U(x)$, można rozwiązać równanie Schrödingera dla tego elektronu i na tej podstawie wyznaczyć interesującą nas zależność dyspersyjną $W(k)$. Przykład takiego postępowania będzie przedmiotem rozważań w następnym punkcie. Obecnie poprzestaniemy na pewnym ogólnym stwierdzeniu, dotyczącym postaci funkcji falowej $\psi(x)$ dla elektronu poruszającego się w jednowymiarowym polu periodycznym.

Ze względu na to, że potencjał pola, w którym porusza się elektron, wykazuje periodyczność z okresem sieci, rozkład gęstości prawdopodobieństwa znalezienia elektronu musi być w każdej z komórek, przedstawionych na rys. 6/9, identyczny. Ponieważ gęstość prawdopodobieństwa znalezienia elektronu w punkcie x jest proporcjonalna do $|\psi(x)|^2$ zatem sformułowany wyżej warunek można wyrazić wzorem

$$|\psi(x+a)|^2 = |\psi(x)|^2 \tag{6/7}$$

Funkcja falowa elektronu w polu periodycznym musi więc mieć taką postać, aby dla każdego x była spełniona równość (6/7). Wypływa stąd wniosek, że funkcja ta musi być wyrażona wzorem

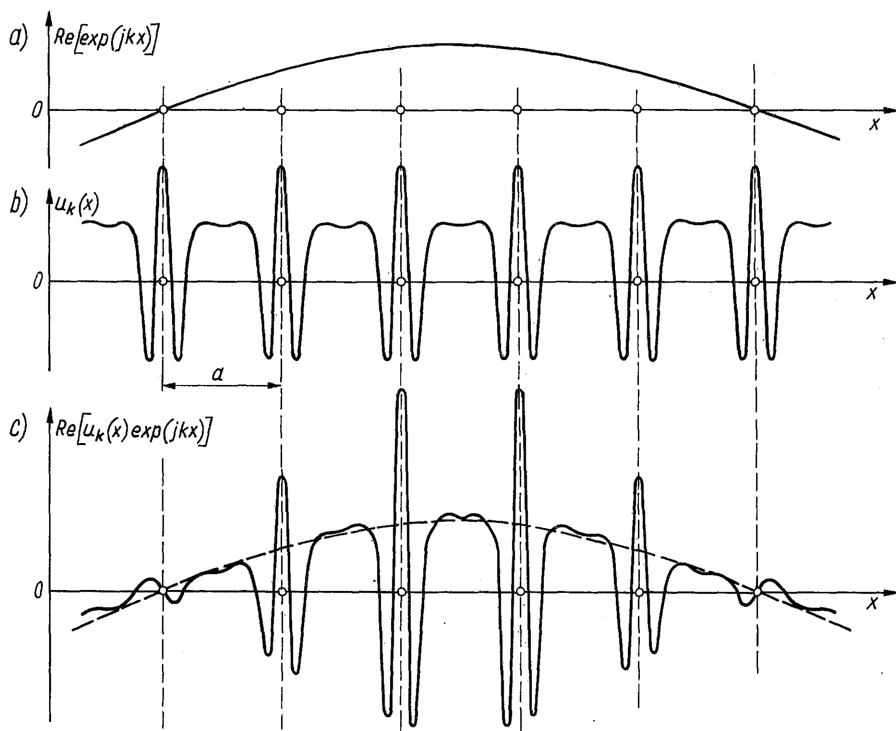
$$\psi(x) = u_k(x)\exp(\mathrm{j}\,kx) \tag{6/8}$$

gdzie: k jest wektorem falowym, który może przyjmować dowolną wartość, $u_k(x)$ zaś – funkcją zespoloną o okresie a. Konkretna postać funkcji $u_k(x)$ zależy od funkcji $U(x)$ podstawionej do równania Schrödingera. Dla każdej wartości wektora k ma ona przy tym inny kształt, dlatego też oznaczamy ją symbolem $u_k(x)$.

Wniosek wyrażony wzorem (6/8) nosi nazwę *twierdzenia Blocha*, a funkcje o postaci (6/8) nazywamy *funkcjami Blocha*. Nie będziemy tu udowodniali tego twierdzenia, łatwo jest natomiast wykazać słuszność twierdzenia odwrotnego. Jeżeli bowiem funkcja falowa ma postać (6/8), to kwadrat jej modułu

$$|\psi(x)|^2 = |u_k(x)|^2 \tag{6/9}$$

Ponieważ funkcja $u_k(x)$ jest periodyczna z okresem a, zatem warunek (6/7) jest spełniony. Rozkład gęstości prawdopodobieństwa znalezienia elektronu wy-



6/10 Przykładowy przebieg jednowymiarowej funkcji Blocha: a) funkcja modulowana; b) funkcja modulująca; c) wypadkowa funkcja falowa

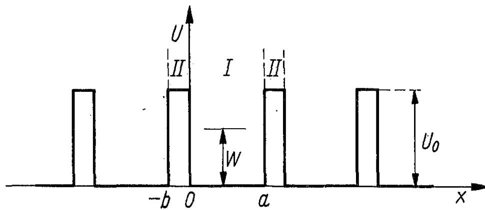
kazuje więc periodyczność z okresem sieci, pomimo tego, że sama funkcja $\psi(x)$ nie jest periodyczna z tym okresem – zob. rys. 6/10.

Analizując wyrażenie (6/8), dostrzegamy, że czynnik exp (j kx) reprezentuje funkcję falową elektronu swobodnego. Funkcja falowa elektronu w polu periodycznym ma więc postać fali płaskiej zmodulowanej czynnikiem $u_k(x)$ o okresowości sieci.

6.3 MODEL KRONIGA-PENNEY'A

Przystąpimy teraz do rozwiązania równania Schrödingera dla elektronu poruszającego się w polu periodycznym. W tym celu przyjmiemy najprostszą postać jednowymiarowego potencjału efektywnego, zaproponowaną przez Kroniga i Penney'a i przedstawioną na rys. 6/11. Potencjał ten odtwarza wprawdzie w niewielkim stopniu warunki występujące w rzeczywistym krysztale, uzyskane rozwiązanie umożliwia jednak sformułowanie pewnych ogólnych wniosków dotyczących zależności dyspersyjnej dla elektronu w ośrodku periodycznym.

W modelu przedstawionym na rys. 6/11 mamy do czynienia z nieskończonym łańcuchem utworzonym ze studni potencjału o szerokości a, oddzielonych od siebie barierami o szerokości b. Potencjał ten wykazuje więc periodyczność



6/11 Rozkład energii potencjalnej w modelu Kroniga-Penney'a

o okresie równym $a+b$. W polu opisanym tym potencjałem porusza się elektron o całkowitej energii W, mniejszej od wysokości bariery U_0.

Rozwiązywaniem takich problemów zajmowaliśmy się w rozdz. 2. Podobnie jak tam, należy więc znaleźć rozwiązania równania Schrödingera niezależnie dla obszaru I i dla obszaru II, a następnie dopasować je do siebie tak, aby uzyskana funkcja falowa miała postać funkcji Blocha i wraz ze swą pochodną spełniała warunki ciągłości na granicach obszarów.

Dla obszaru I, w którym $U = 0$, rozwiązanie równania Schrödingera ma postać

$$\psi_1(x) = A\exp(\mathrm{j}\,\alpha x) + B\exp(-\mathrm{j}\,\alpha x) \tag{6/10}$$

gdzie

$$\alpha = \sqrt{\frac{2mW}{\hbar^2}} \tag{6/11}$$

natomiast dla obszaru II, w którym $U = U_0$

$$\psi_2(x) = C\exp(\beta x) + D\exp(-\beta x) \tag{6/12}$$

przy czym

$$\beta = \sqrt{\frac{2m(U_0 - W)}{\hbar^2}} \tag{6/13}$$

Opierając się na wzorze (6/8), można wyznaczyć kształt funkcji $u(x)$ w obszarach I i II.

$$u_1(x) = A\exp[\mathrm{j}(\alpha - k)x] + B\exp[-\mathrm{j}(\alpha + k)x] \tag{6/14}$$

$$u_2(x) = C\exp[(\beta - \mathrm{j}k)x] + D\exp[-(\beta + \mathrm{j}k)x] \tag{6/15}$$

W celu wyznaczenia współczynników A, B, C, D skorzystamy z warunku, że funkcja $u(x)$ i jej pochodna muszą być ciągłe na granicy obszarów I i II, tzn. w punkcie $x = 0$, oraz że $u(x)$ jest funkcją okresową o okresie $a+b$. Funkcje $u_1(x)$ i $u_2(x)$ muszą więc spełniać następujące warunki:

$$\left.\begin{aligned} u_1(0) &= u_2(0) \\ \left(\frac{\mathrm{d}u_1}{\mathrm{d}x}\right)_{x=0} &= \left(\frac{\mathrm{d}u_2}{\mathrm{d}x}\right)_{x=0} \\ u_1(a) &= u_2(-b) \\ \left(\frac{\mathrm{d}u_1}{\mathrm{d}x}\right)_{x=a} &= \left(\frac{\mathrm{d}u_2}{\mathrm{d}x}\right)_{x=-b} \end{aligned}\right\} \tag{6/16}$$

Podstawiając do tych równań wyrażenia (6/14) i (6/15), otrzymuje się układ 4 równań algebraicznych jednorodnych, w którym niewiadomymi są współczynniki A, B, C, D. Układ ten ma niezerowe rozwiązanie tylko wówczas, gdy jego wyznacznik jest równy zeru. Wymaganie to prowadzi do następującej zależności:

$$\frac{\beta^2-\alpha^2}{2\alpha\beta}\,\mathrm{sh}\,\beta b \sin\alpha a+\mathrm{ch}\,\beta b\cos\alpha a = \cos k\,(a+b) \tag{6/17}$$

Otrzymane równanie umożliwia wyznaczenie wartości współczynników α i β, a tym samym również energii W dla zadanej wartości wektora falowego k. W równaniu (6/17) tkwi więc poszukiwana zależność dyspersyjna $W(k)$.

Równanie to może być rozwiązane metodą graficzną lub numeryczną. W celu dalszego uproszczenia zagadnienia zawęzimy nasze rozważania do przypadku granicznego, w którym $U_0 \to \infty$, natomiast $b \to 0$ i to w taki sposób, że iloczyn bU_0 zachowuje stałą wartość. Równanie (6/17) sprowadza się wówczas do postaci

$$Q\,\frac{\sin\alpha a}{\alpha a}+\cos\alpha a = \cos ka \tag{6/18}$$

gdzie

$$Q = \frac{ma}{\hbar^2}\,bU_0 \tag{6/19}$$

Wprowadzony tu współczynnik Q, proporcjonalny do iloczynu bU_0, można traktować jako miarę trudności pokonania przez elektron bariery o szerokości b i wysokości U_0.

Zanim przystąpimy do zapoznania się z rozwiązaniem równania (6/18), rozpatrzmy następujące 2 przypadki graniczne:

a) jeżeli $Q \to 0$ (tzn. bariery utrzymują skończoną wysokość lecz stają się nieskończenie cienkie), to

$$\cos ka = \cos\alpha a \tag{6/20}$$

a zatem

$$k = \pm\alpha = \pm\sqrt{\frac{2mW}{\hbar^2}} \tag{6/21}$$

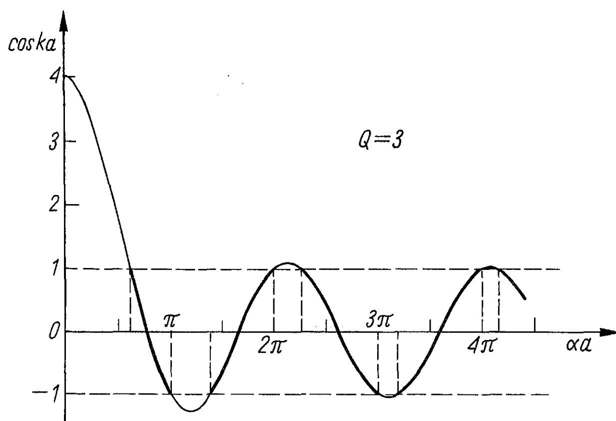
Elektron zachowuje się wówczas tak, jak elektron swobodny, a zależność dyspersyjna jest identyczna z podaną wzorem (2/13).

b) jeżeli $Q \to \infty$ (tzn. bariery utrzymują skończoną szerokość, a ich wysokość wzrasta nieograniczenie), to zależność (6/18) przyjmuje postać

$$\sin\alpha a = 0 \tag{6/22}$$

Jest to równanie analogiczne do wzoru (2/51), który otrzymaliśmy rozpatrując zachowanie się elektronu w studni potencjału o nieskończenie wysokich ściankach. W analizowanym tu przypadku elektron jest więc uwięziony w studni, a jego energia może przyjmować jedną z wartości własnych określonych zależnością (6/22).

Dla skończonych wartości parametru Q równanie (6/18) może być rozwiązane metodą graficzną w sposób przedstawiony na rys. 6/12. Na osi rzędnych odłożono tu $\cos ka$, a więc wielkość zależną od wektora falowego rozpatrywanego elektronu, na osi odciętych zaś – iloczyn αa, czyli wielkość zależną od energii elektronu. Ponieważ $|\cos ka|$ może w dziedzinie liczb rzeczywistych przyjmować tylko wartości nie większe niż jeden, zatem dopuszczalne są tylko takie wartości αa (i odpowiadające im wartości energii), którym odpowiadają punkty krzywej nie wykraczające poza proste graniczne $\cos ka = \pm 1$. Te części krzywej zaznaczono na rys. 6/12 linią pogrubioną.



6/12 Przykładowy przebieg funkcji cos *ka* wyrażonej wzorem (6/18)

Najważniejszym wnioskiem, jaki można wyciągnąć ze wzoru (6/18) i ze sporządzonego na jego podstawie rys. 6/12, jest to, że dopuszczalne wartości energii elektronu poruszającego się w polu periodycznym tworzą pewne pasma, przedzielone obszarami energii zabronionych (przerwami energetycznymi). W tych obszarach zabronionych $|\cos ka|$ przyjmuje wartości większe od jedności, co oznacza, że wektor falowy k staje się urojony. Podstawiając $k = -jk'$, można wówczas funkcję falową wyrazić wzorem

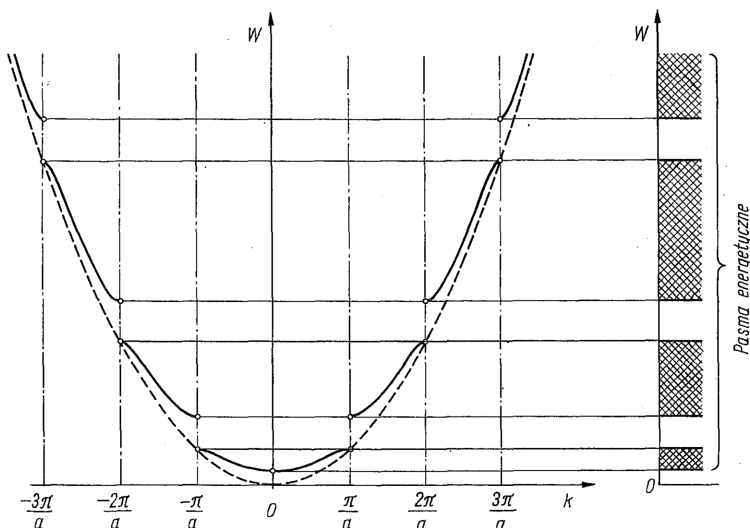
$$\psi(x) = u_k(x) \exp(-k'x) \tag{6/23}$$

Fala taka nie może rozchodzić się w krysztale. Jest to fala zanikająca, podobnie jak fala występująca w obszarze klasycznie zabronionym w przypadku rozpatrywanym w p. 2.4.

Na podstawie wzoru (6/18) można sporządzić wykres przedstawiający zależność $W(k)$ dla rozpatrywanego modelu kryształu[1]. Wykres taki przedstawiono na rys. 6/13. Linią przerywaną pokazano tam zależność dyspersyjną dla elektronu swobodnego. Zależności te różnią się wskutek tego, że elektron w kry-

[1] Sporządzenie takiego wykresu nie jest problemem jednoznacznie określonym, ponieważ danej wartości $\cos ka$ odpowiada nieskończenie wiele wartości wektora falowego k. Zagadnienie tej niejednoznaczności będzie omawiane w następnym punkcie. Chwilowo przyjmiemy, że każdej wartości $\cos ka$ odpowiadają tylko 2 różniące się znakiem wartości k.

sztale oddziałuje z siecią krystaliczną, tzn. ze wszystkimi pozostałymi cząstkami układu – jądrami atomowymi i elektronami. Przestrzennie periodyczny charakter tych oddziaływań jest odpowiedzialny za podział całego zakresu energii na pasma i przerwy energetyczne.



6/13 Zależność dyspersyjna $W(k)$ dla elektronu w modelu Kroniga-Penney'a

Jak widać na rys. 6/13, wewnątrz każdego pasma energetycznego funkcja $W(k)$ wzrasta monotonicznie, a na brzegach pasm występują ekstrema i nieciągłości tej funkcji. Nieciągłości te pojawiają się przy takich wartościach wektora falowego, przy których $|\cos ka| = 1$, a więc

$$|k| = \frac{n\pi}{a} \tag{6/24}$$

gdzie n – dowolna liczba naturalna. Podstawiając $k = 2\pi/\lambda$, można warunek ten zapisać w postaci

$$\boxed{n\lambda = 2a} \tag{6/25}$$

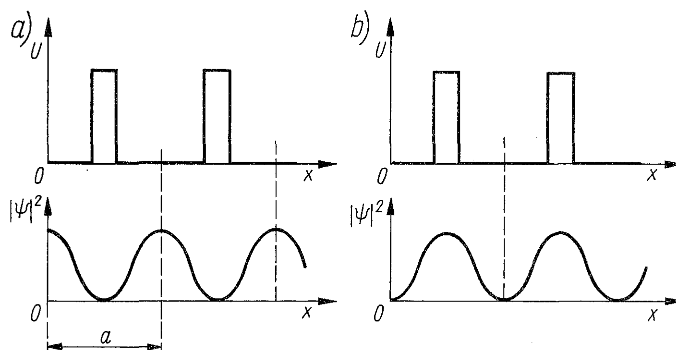
Jest to, wyrażony wzorem (1/30), warunek odbicia braggowskiego dla fali o kącie ϑ równym 90°. Wynika stąd, że *przerwy energetyczne dla elektronu w krysztale występują wówczas, gdy jest spełniony warunek braggowskiego odbicia fal elektronowych od płaszczyzn sieciowych kryształu.* Elektron o długości fali wyrażonej wzorem (6/25) nie może więc poruszać się w krysztale.

Posługując się wzorem (6/5), można wyznaczyć prędkość elektronu (czyli prędkość grupową fali elektronowej) w krysztale. Na brzegach pasm stwierdzamy, że

$$\frac{dW}{dk} = 0 \tag{6/26}$$

a zatem prędkość grupowa jest równa zeru. Fala elektronowa jest w tym przypadku falą stojącą, której nie towarzyszy transport energii. Fala ta powstaje w wyniku interferencji fali padającej i fali odbitej od płaszczyzn sieciowych kryształu.

W zależności od faz interferujących ze sobą fal możemy mieć do czynienia z jedną z 2 sytuacji przedstawionych na rys. 6/14. W pierwszym wypadku (rys. 6/14a) maksima gęstości prawdopodobieństwa znalezienia elektronu przypadają na te obszary kryształu, w których występują minima energii potencjalnej. Sytuacji tej odpowiada zatem mniejsza wartość energii elektronu niż w drugim przypadku (rys. 5/14b), gdy największe prawdopodobieństwo znalezienia elektronu występuje w obszarach, w których energia potencjalna ma dużą wartość. W ten sposób można wyjaśnić fakt, że na brzegach pasm każdej wartości wektora falowego (z wyjątkiem $k = 0$) odpowiadają 2 różne stany elektronowe, tzn. 2 rozwiązania równania Schrödingera różniące się wartością energii elektronu.

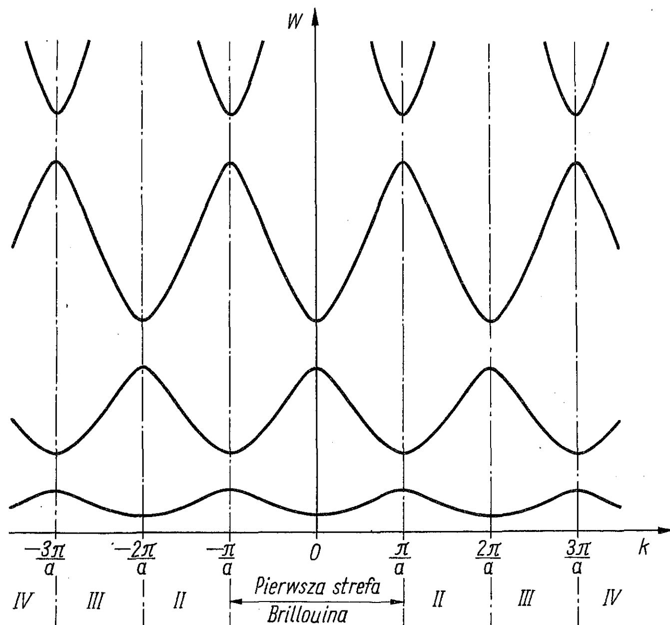


6/14 Rozkłady gęstości prawdopodobieństwa znalezienia elektronu w modelu Kroniga-Penney'a ($k = \pi/a$): a) stan o mniejszej energii); b) stan o większej energii

Dla elektronu swobodnego, dla którego zachodzi kwadratowa zależność energii od wartości wektora falowego, wektor ten jest ściśle związany z pędem elektronu – wzór (1/37). Dla elektronu w krysztale energia nie jest, jak to widać na rys. 6/13, proporcjonalna do k^2, dlatego też wektora $\hbar \boldsymbol{k}$ nie można utożsamiać z pędem elektronu. Wielkość ta ma jednak, jak dalej zobaczymy, pewne cechy pędu, w związku z czym nosi ona nazwę *pseudopędu* lub *pędu krystalicznego*. W książce tej nie będziemy jednak czynili tego rozróżnienia i będziemy posługiwali się nazwą „pęd" bez dodatku „krystaliczny".

Jak już wspomniano, każdej wartości energii elektronu odpowiada nieskończenie wiele wartości wektora falowego. Spośród nich uwzględniono na rys. 6/13 tylko po jednej parze różniącej się znakiem, tzn. odpowiadającej przeciwnym kierunkom ruchu elektronu. Pełny wykres uwzględniający wszystkie wartości wektora $\boldsymbol{k}$ (w obranym przedziale) przedstawiono na rys. 6/15.

Jeżeli weźmiemy pod uwagę tylko jeden kierunek ruchu elektronu, to, jak wynika z rys. 6/15, dowolnie obranej wartości energii elektronu należy przypo-



6/15 Pełny wykres zależności W (k) dla modelu Kroniga-Penney'a

rządkować następujący zbiór wartości wektora falowego:

$$k_n = k_0 + \frac{2\pi n}{a} \tag{6/27}$$

gdzie: $n = 0, \pm 1, \pm 2, \ldots$

Jednocześnie możemy zauważyć, że danej wartości wektora falowego odpowiada wiele różnych wartości energii – po jednej w każdym pasmie. Energia jest więc wieloznaczną funkcją wektora k.

6.4 OGÓLNE WŁASNOŚCI PASM ENERGETYCZNYCH

Z rysuku 6/15 można łatwo odczytać kilka własności funkcji $W(k)$. Ponieważ jednak własności te dotyczą nie tylko modelu Kroniga-Penney'a, ale mają ogólniejszy charakter, omówieniu ich poświęcamy odrębny punkt.

Zajmiemy się tu jednowymiarowym modelem kryształu o dowolnym rozkładzie potencjału efektywnego, tzn. dowolnym kształcie funkcji $U(x)$. Zakładamy jedynie, że potencjał ten jest periodyczny z okresem równym stałej sieci a.

Rozpatrywany model ma bezpośredni związek z rzeczywistymi trójwymiarowymi strukturami krystalicznymi, ponieważ można go uważać za reprezentację zbioru płaszczyzn sieciowych prostopadłych do danego kierunku. Wnioski uzyskane z analizy modelu jednowymiarowego mają więc zastosowanie również do stanów elektronowych związanych z wybranym kierunkiem w sieci trójwymiarowej.

Omawiając własności funkcji $W(k)$ należy przede wszystkim stwierdzić, że jest to funkcja analityczna – jest więc ciągła, różniczkowalna i może być rozwijana w szereg potęgowy. Dalej, charakteryzuje się ona parzystością, tzn.

$$W(-k) = W(k) \tag{6/28}$$

Wynika to stąd, że zmieniając kierunek ruchu elektronu na przeciwny, otrzymuje się rozwiązanie fizycznie równoważne. Istotnie, przy odwróceniu kierunku ruchu elektronu rozwiązanie równania Schrödingera przyjmuje postać

$$\psi(x) = u_k^*(x)\exp(-jkx) \tag{6/29}$$

gdzie $u_k^*(x)$ – funkcja sprzężona z $u_k(x)$. Funkcji tej odpowiada identyczny rozkład gęstości prawdopodobieństwa znalezienia elektronu, a zatem również identyczna wartość energii jak w przypadku opisanym funkcją (6/8).

Bezpośrednią konsekwencją symetrii translacyjnej kryształu jest periodyczność zależności dyspersyjnej $W(k)$. Wynika to z faktu, że stany elektronowe o wektorach falowych związanych zależnością (6/27) są sobie pod każdym względem fizycznie równoważne. Możemy się o tym przekonać, przekształcając zależność (6/8) do następującej postaci:

$$\psi(x) = \left[u_k(x)\exp\left(-j\frac{2\pi n}{a}x\right)\right]\exp\left[j\left(k+\frac{2\pi n}{a}\right)x\right] \tag{6/30}$$

Ponieważ czynnik $\exp\left(-j\frac{2\pi n}{a}x\right)$ jest periodyczny z okresem a, zatem całe wyrażenie

$$u_k(x)\exp\left(-j\frac{2\pi n}{a}x\right)$$

może być traktowane jako funkcja modulująca, odpowiadająca stanowi o wartości wektora falowego wyrażonej wzorem

$$k' = k+\frac{2\pi n}{a} \tag{6/31}$$

Możemy więc napisać, że

$$\psi(x) = u_{k+\frac{2\pi n}{a}}(x)\exp\left[j\left(k+\frac{2\pi n}{a}\right)x\right] \tag{6/32}$$

Jeżeli funkcja falowa o postaci $\psi(x) = u_k(x)\exp(jkx)$ jest rozwiązaniem równania Schrödingera, to także funkcja (6/32) spełnia to równanie. Obu tym stanom odpowiada taki sam rozkład gęstości prawdopodobieństwa znalezienia elektronu, ponieważ

$$\left|u_k(x)\exp\left(-j\frac{2\pi n}{a}x\right)\right| = |u_k(x)| \tag{6/33}$$

W obu stanach energia elektronu ma więc tę samą wartość. Funkcja $W(k)$ spełnia zatem warunek

$$W\left(k+\frac{2\pi n}{a}\right) = W(k) \tag{6/34}$$

dla dowolnej liczby całkowitej n. Jak wynika z tej równości, okres funkcji $W(k)$ jest równy $2\pi/a$. Wielkość $1/a$ nazywamy *stałą sieci odwrotnej*[1].

Jeżeli w punkcie $k = 0$ nie występuje degeneracja, to z parzystości funkcji $W(k)$ wynika zerowanie się w tym punkcie jej pochodnej. W punkcie $k = 0$ występuje zatem ekstremum funkcji. Identyczne ekstrema pojawiają się oczywiście również w punktach $k = \pm 2\pi n/a$. Najprostszym (a dla sieci jednowymiarowej jedynym możliwym) przebiegiem $W(k)$ mającym wyszczególnione własności jest przebieg o charakterze przedstawionym na rys. 6/15.

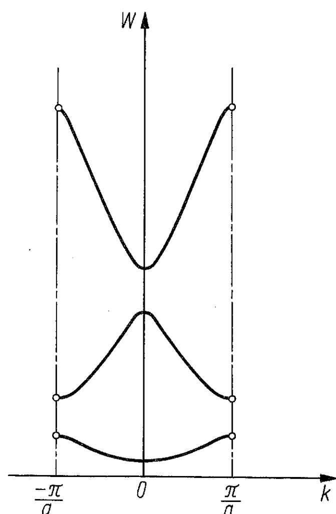
Periodyczność tej zależności umożliwia ograniczenie przedziału zmienności wektora k do pewnego obszaru, który dla sieci jednowymiarowej jest odcinkiem o długości $2\pi/a$. Środek tego odcinka jest dogodnie obrać w punkcie $k = 0$. Tak zdefiniowany obszar nosi nazwę *pierwszej strefy Brillouina*[2]. Numerację dalszych stref podano na rys. 6/15.

Wykres $W(k)$ ograniczony do pierwszej strefy Brillouina pokazano na rys. 6/16. Wektory falowe należące do tej strefy, tzn. spełniające warunki

$$-\frac{\pi}{a} < k \leqslant \frac{\pi}{a} \qquad (6/35)$$

nazywamy *zredukowanymi* wektorami falowymi. W dalszych rozważaniach będziemy się posługiwali tylko takimi wektorami.

Na rysunku 6/16 punkty $k = -\frac{\pi}{a}$ i $k = +\frac{\pi}{a}$ są równoważne. Jeżeli więc wektor falowy zmienia się w sposób ciągły, np. wzrasta, co, jak zobaczymy w następnym punkcie, zachodzi pod działaniem siły zewnętrznej, to w pewnej



6/16
Fragment wykresu zależności $W(k)$ ograniczony do pierwszej strefy Brillouina

[1] W książce tej nie będziemy posługiwali się pojęciem sieci odwrotnej. Zainteresowany Czytelnik może znaleźć informacje na ten temat w każdym podręczniku z zakresu fizyki ciała stałego.

[2] Czytaj: brijuę.

chwili dochodzi on do granicy strefy ($k = \pi/a$). Na wykresie przedstawionym na rys. 6/15 można dalej przesuwać się wzdłuż osi k, ale na mocy periodyczności funkcji $W(k)$ jest to równoważne przeniesieniu się elektronu do punktu $k = -\pi/a$. Tym sposobem elektron kontynuuje swój ruch w obrębie pierwszej strefy Brillouina.

6.5 DYNAMIKA ELEKTRONÓW W CIELE STAŁYM

Zastanówmy się teraz, jak będzie się zachowywał elektron, tzn. jak będzie się zmieniał jego stan kwantowy, pod wpływem siły zewnętrznej. Dla uproszczenia przyjmijmy, że w rozpatrywanym pasmie energetycznym znajduje się tylko jeden, rozpatrywany przez nas elektron.

Jeżeli elektron porusza się z prędkością v, to w wyniku działania siły F jego energia wzrośnie w czasie dt o wartość $dW = Fv\,dt$. Stąd

$$\frac{dW}{dt} = Fv \tag{6/36}$$

Prędkość elektronu w ciele stałym może być wyrażona wzorem (6/5). Wzór ten został wprawdzie wyprowadzony dla elektronu swobodnego, stosuje się on jednakże ogólnie, ponieważ zgodnie z zasadami mechaniki kwantowej poruszającej się cząstce przypisujemy prędkość grupową zdefiniowaną wzorem (1/23).

Korzystając ze wzoru (6/5), możemy lewą stronę wyrażenia (6/36) przekształcić w następujący sposób:

$$\frac{dW}{dt} = \frac{dW}{dk}\,\frac{dk}{dt} = \hbar v\,\frac{dk}{dt} \tag{6/37}$$

Ze wzorów (6/36) i (6/37) otrzymuje się równanie ruchu elektronu w jednowymiarowym polu periodycznym

$$\boxed{\hbar\,\frac{dk}{dt} = F} \tag{6/38}$$

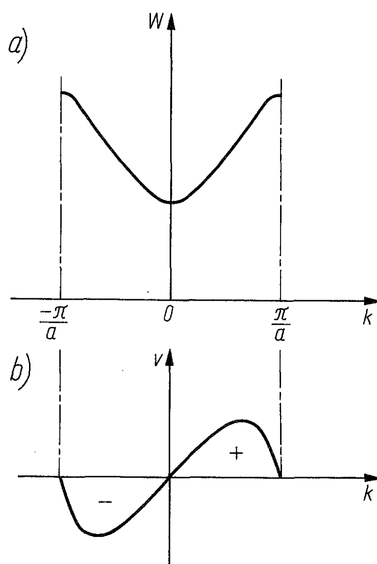
Lewa strona tego równania przedstawia pochodną pędu względem czasu. Otrzymany wzór ma więc postać analogiczną do równania wyrażającego drugą zasadę dynamiki Newtona. Należy jednak zauważyć, że w równaniu Newtona występuje suma wszystkich sił działających na dane ciało, tu zaś występuje tylko siła *zewnętrzna*. Na elektron w ciele stałym działają, jak wiemy, także siły wewnętrzne, związane z oddziaływaniem wszystkich pozostałych elementów kryształu. Wpływ tych sił został już wcześniej uwzględniony przez nadanie funkcji falowej elektronu postaci funkcji Blocha o odpowiednio dobranym czynniku modulującym $u_k(x)$ zależnym od rozkładu przestrzennego tych sił.

Jeżeli na elektron w krysztale nie działa żadna siła zewnętrzna, to, jak wynika ze wzoru (6/38), wektor falowy ma określoną, niezmienną w czasie wartość. Wartości tej odpowiada pewien stan stacjonarny o określonej energii. Oznacza to,

że elektron porusza się w krysztale ruchem jednostajnym, nie doznając żadnych zakłóceń ze strony sieci krystalicznej. Średnia droga swobodna elektronu jest w tych warunkach nieskończenie wielka. Jeżeli w rzeczywistości ma ona wartość skończoną, to jest to spowodowane odchyleniami rzeczywistej struktury krystalicznej od idealnej periodyczności. Odchylenia te, wywołane drganiami cieplnymi rdzeni atomowych oraz defektami strukturalnymi powodują rozpraszanie fali elektronowej. W wyniku tych rozproszeń wektor falowy zmienia swą wartość, czemu towarzyszy zmiana prędkości i innych parametrów ruchu elektronu.

Przy oddziaływaniach elektronu w ciele stałym z innymi cząstkami, np. innymi elektronami, fotonami lub fononami, obowiązuje, podobnie jak dla cząstek swobodnych, zasada zachowania pędu. Ponieważ jednak wektor falowy jest określony z dokładnością do stałej addytywnej $2\pi/a$, zatem zasada zachowania pędu w ciele stałym musi być sformułowana w następujący sposób: *suma pędów oddziałujących ze sobą cząstek „przed zderzeniem" jest równa sumie tych pędów „po zderzeniu" z dokładnością do dowolnej całkowitej wielokrotności stałej h/a.*

Przejdźmy teraz do przypadku, w którym na elektron działa stała w czasie siła F. Jak wynika ze wzoru (6/38), wektor k będzie się wówczas zmieniał liniowo w czasie. Jeżeli przyjmiemy, że w chwili początkowej mieliśmy $k = 0$, to, jak widzimy na rys 6/17, prędkość elektronu będzie wzrastała od zera do pewnej wartości maksymalnej, a następnie malała, przy czym osiągnie ona ponownie wartość równą zeru przy $k = \pi/a$. Punkt ten, jak to już stwierdziliśmy, jest równoważny punktowi $k = -\pi/a$. Elektron „przeskoczy" więc jak gdyby na drugi brzeg strefy Brillouina i będzie kontynuował swój ruch. Prędkość przyjmuje teraz wartości ujemne, a zatem ruch odbywa się w kierunku przeciwnym do przyłożonej siły. Po upływie pewnego czasu elektron zatrzyma się w punkcie $k = 0$, po czym cały proces rozpocznie się od początku.



6/17
Pasmo przewodnictwa półprzewodnika (a) i przebieg zależności prędkości elektronu od wektora falowego (b)

W prawie pustym pasmie energetycznym elektron poddany działaniu stałej w czasie siły wykonywałby więc ruch drgający, a jego energia zmieniałaby się od pewnej wartości minimalnej na dnie pasma do wartości maksymalnej na jego wierzchołku. Bezpośrednie zaobserwowanie takiego ruchu elektronu jest niemożliwe, ponieważ w każdym rzeczywistym krysztale fala elektronowa ulega rozpraszaniu na zakłóceniach regularności sieci krystalicznej.

Na podstawie wzorów (6/5) i (6/38) można wyznaczyć przyspieszenie elektronu w ciele stałym. Wynosi ono

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{1}{\hbar^2} \frac{\mathrm{d}^2 W}{\mathrm{d}k^2} F \tag{6/39}$$

Wzór ten przyjmuje postać analogiczną do równania ruchu Newtona, jeżeli formalnie wprowadzi się następujący współczynnik proporcjonalności między przyspieszeniem a siłą

$$\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{\mathrm{d}^2 W}{\mathrm{d}k^2} \tag{6/40}$$

Dla elektronu swobodnego współczynnikiem tym jest odwrotność masy, którą można wyznaczyć z zależności dyspersyjnej $W(k)$ – zob. wzór (6/6). Z tego względu wielkość m^* zdefiniowana wzorem (6/40) nazywamy *masą efektywną* elektronu w ciele stałym.

Parametr ten ma dość specyficzne własności. Po pierwsze różni się on na ogół, i to niekiedy bardzo znacznie, od masy elektronu swobodnego m_0[1]. Dalej masa efektywna zależy od wartości energii elektronu, przyjmuje więc różne wartości (dla tej samej wartości wektora falowego) w różnych pasmach energetycznych. W obrębie danego pasma zmienia się wraz z wartością wektora k, może przy tym przyjmować zarówno wartości dodatnie, jak i ujemne, a w punkcie przegięcia krzywej $W(k)$ staje się nieskończenie wielka. Na tym nie kończą się jeszcze osobliwości tego nowego pojęcia, z którym zapoznaliśmy się. Dla rozpatrywanej tu sieci jednowymiarowej wielkość ta jest skalarem, przy czym ze względu na parzystość funkcji $W(k)$ nie zależy ona od kierunku ruchu elektronu, natomiast dla sieci trójwymiarowej wyrażenie $1/m^*$ jest tensorem, o czym będzie mowa w p. 7.1.

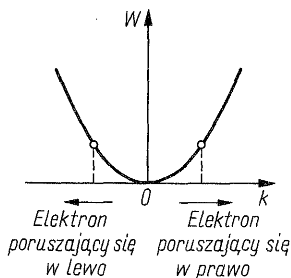
Z rozważań zawartych w tym punkcie można jeszcze wyciągnąć następujący ważny wniosek: dla pasma całkowicie zapełnionego średnia prędkość elektronów, zdefiniowana wzorem

$$\langle v \rangle = \sum_k v(k) \tag{6/41}$$

jest równa zeru (sumowanie jest tu rozciągnięte na wszystkie wartości zredukowanego wektora falowego). Z parzystości funkcji $W(k)$ wynika bowiem, że dla

---

[1] Spotyka się na przykład półprzewodniki, w których masa efektywna elektronów wynosi kilka setnych m_0, schodząc w niektórych przypadkach – nawet do $3 \cdot 10^{-4}$ m_0.



6/18
Fragment pasma przewodnictwa w otoczeniu punktu $k = 0$

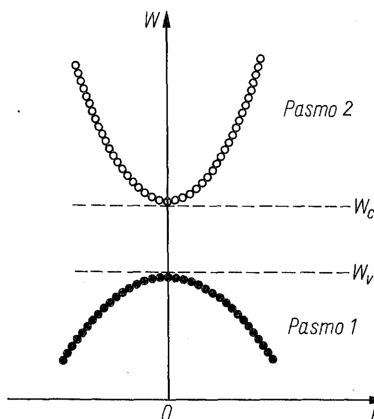
każdego elektronu o prędkości $v(k)$ istnieje elektron o przeciwnie skierowanym wektorze falowym i przeciwnym kierunku prędkości – zob. rys. 6/18. Tym tłumaczy się omawiany w p. 6.1.1 fakt, że prąd elektryczny wytwarzany przez wszystkie elektrony znajdujące się w pasmie całkowicie zapełnionym jest równy zeru.

6.6 PASMA PARABOLICZNE

Rozpatrzmy teraz pewien fragment rys. 6/16 i załóżmy, że najwyższe pokazane tam pasmo jest pasmem przewodnictwa półprzewodnika. Pasmo znajdujące się bezpośrednio pod nim jest zatem pasmem walencyjnym. Niższe pasma nie będą nas interesowały, ponieważ w normalnie spotykanych warunkach są one całkowicie zapełnione, nie mają zatem wpływu na własności elektryczne i optyczne półprzewodnika.

W temperaturze zera bezwzględnego pasmo walencyjne, jak już wyżej mówiliśmy, jest całkowicie obsadzone przez elektrony, a pasmo przewodnictwa – puste. Pokazano to w sposób poglądowy na rys. 6/19.

W temperaturach wyższych od 0 K (np. w temperaturze pokojowej) część elektronów z pasma walencyjnego zostaje wzbudzona do pasma przewodnictwa. Elektrony te obsadzają poziomy znajdujące się w pobliżu dna tego pasma. Jak zobaczymy w rozdz. 8, energie tych elektronów (liczone od dna pasma) nie prze-



6/19
Obsadzenie pasma przewodnictwa i pasma walencyjnego półprzewodnika w temperaturze zera bezwzględnego

kraczają, praktycznie biorąc, wartości rzędu 6 kT, która w temperaturze pokojowej wynosi ok. 150 meV. Z tego względu w pasmie przewodnictwa interesuje nas prawie wyłącznie jego dolny odcinek. Rozwijając funkcję $W(k)$ w szereg potęgowy i uwzględniając jej parzystość, można dla interesującego nas odcinka tego pasma napisać

$$W(k) = W_c + \left(\frac{d^2 W}{dk^2}\right)_{k=0} k^2 + \dots \tag{6/42}$$

Jeżeli ograniczymy to rozwinięcie do 2 pierwszych wyrazów, to zależność energii od wektora falowego przyjmie postać formy kwadratowej. Pasma, dla których przybliżenie takie jest dopuszczalne, nazywamy *pasmami parabolicznymi*. Dla takich pasm można zależność (6/42) przedstawić w następującej postaci, analogicznej do odpowiedniej zależności dla elektronu swobodnego:

$$W(k) = W_c + \frac{\hbar^2 k^2}{2m_2^*} \tag{6/43}$$

gdzie m_2^* – masa efektywna elektronu w pasmie przewodnictwa (które na rysunku oznaczyliśmy jako pasmo *2*). Jest ona zdefiniowana wzorem (6/40) i, jak widać, jest dodatnia. Zgodnie z założeniem paraboliczności pasma masa efektywna jest wielkością stałą, niezależną od energii elektronu[1].

Z kolei w pasmie walencyjnym elektrony obsadzają całkowicie niższe poziomy energetyczne, pozostawiając częściowo nieobsadzone poziomy znajdujące się w pobliżu wierzchołka pasma. Postępując analogicznie, jak w przypadku pasma przewodnictwa, otrzymamy dla pasma parabolicznego następujący wzór:

$$W(k) = W_v + \frac{\hbar^2 k^2}{2m_1^*} \tag{6/44}$$

gdzie m_1^* – masa efektywna elektronu w pasmie walencyjnym. Jest ona zdefiniowana tym samym wzorem (6/40), ale ze względu na inny kształt funkcji $W(k)$ ma wartość ujemną, $|m_1|$ różni się przy tym z reguły od m_2.

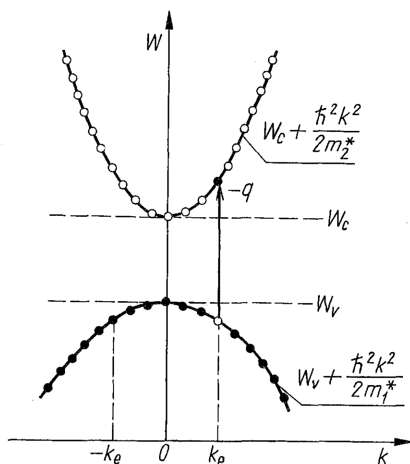
Problem 6.2. Wykaż, że dla dowolnego pasma parabolicznego zachodzi następujący związek:

$$m^* v = \hbar k \tag{6/45}$$

6.7 POJĘCIE I WŁASNOŚCI DZIURY

Omawiane w poprzednim punkcie fragmenty pasm przedstawiono na rys. 6/20. Pokazano tu w sposób poglądowy poszczególne stany kwantowe. Zastanówmy się

[1] Dla pasm nieparabolicznych masa efektywna, jak to już wyżej zaznaczono, jest zależna od energii elektronu. Sprawa komplikuje się ponadto ze względu na to, że w różnych okolicznościach występują różnie zdefiniowane masy efektywne. W związku z tym rozróżnia się na przykład *energetyczną* masę efektywną i *pędową* masę efektywną. Zagadnienia te wykraczają jednak poza ramy tej książki, dlatego też rozważania zawarte w tym punkcie zostały ograniczone do pasm parabolicznych, w których rozróżnienie takie nie jest potrzebne.



6/20
Wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa w półprzewodniku

teraz nad następującą sytuacją: załóżmy, że pod wpływem jakiegoś czynnika zewnętrznego (np. fotonu) 1 elektron z całkowicie zapełnionego pasma walencyjnego został wzbudzony do pasma przewodnictwa, przy czym wartość jego wektora falowego nie uległa zmianie. Przejście takie, przedstawione na rys. 6/20 za pomocą pionowej strzałki, nosi nazwę *przejścia prostego*[1].

W wyniku wzbudzenia elektronu w pasmie walencyjnym oswobodził się 1 stan kwantowy, a w pasmie przewodnictwa został obsadzony jeden stan, jak to pokazano na rys. 6/20.

Z rozważań z poprzedniego punktu wiemy już, jak zachowuje się elektron w prawie pustym pasmie przewodnictwa. Stwierdziliśmy tam, że o własnościach dynamicznych tego elektronu decyduje jego masa efektywna, której wartość zależy od kształtu funkcji $W(k)$. Zanim przejdziemy do analizy sytuacji występującej w pasmie walencyjnym, prześledźmy dalsze losy elektronu wzbudzonego do pasma przewodnictwa. Ma on pewną energię kinetyczną, porusza się bowiem z pewną prędkością, która, jak wiemy, jest proporcjonalna do pochodnej funkcji $W(k)$. W wyniku zderzeń z niedoskonałościami sieci krystalicznej rozpatrywany elektron będzie wytracał swą prędkość, przechodząc na coraz niżej położone poziomy energetyczne. Wskutek tego znajdzie się on w punkcie, w którym $k = 0$, a energia osiąga wartość minimalną[2].

Przejdźmy teraz do pasma walencyjnego. Rozważając zachowanie się elektronów w pasmie prawie całkowicie zapełnionym, mamy do rozwiązania problem, który wydaje się znacznie bardziej skomplikowany, ponieważ liczba elektronów w tym pasmie jest niewyobrażalnie duża. Na domiar złego, część z nich ma masy efektywne ujemne, część – dodatnie, a dla niektórych jest ona nieskończenie

[1] Przejścia, przy których zachodzi istotna zmiana wektora falowego, nazywamy *przejściami skośnymi*.

[2] W opisywanym eksperymencie myślowym wykluczamy możliwość powrotu wzbudzanego elektronu do pasma walencyjnego.

wielka[1]. Okazuje się jednak, że pasmo, w którym wszystkie stany kwantowe z wyjątkiem jednego są obsadzone przez elektrony, można traktować jako obsadzone przez pewną quasi-cząstkę, którą nazywamy *dziurą*. Tej niby-cząstce przypisujemy określony ładunek elektryczny, masę efektywną, energię i inne parametry dynamiczne.

W celu wyznaczenia tych parametrów załóżmy, że liczba wszystkich stanów kwantowych w pasmie walencyjnym rozpatrywanego kryształu wynosi N. Spośród nich N-1 stanów jest obsadzonych przez elektrony. Załóżmy dalej, że elektron wzbudzony do pasma przewodnictwa miał przed wzbudzeniem następujące wartości parametrów dynamicznych: masę efektywną $m_1^* < 0$, wektor falowy k_e, prędkość v_e i energię W_e.

Jeżeli dziura ma reprezentować cały zespół elektronów w prawie całkowicie zapełnionym pasmie, to musimy przypisać jej pęd równy sumie pędów tych elektronów. Oznaczając wektor falowy dziury symbolem k_d, możemy więc napisać

$$\hbar k_d = \sum_{i=1}^{N-1} \hbar k_i \tag{6/46}$$

przy czym sumowanie to jest rozciągnięte na wszystkie N-1 elektronów w rozważanym pasmie. Wzór (6/46) można przekształcić do postaci

$$\hbar k_d = \sum_{i=1}^{N} \hbar k_i - \hbar k_e \tag{6/47}$$

Do sumy włączyliśmy tu pęd brakującego elektronu, a następnie od całości odjęliśmy ten pęd.

Pierwszy składnik w tym wzorze jest równy zeru, ponieważ z parzystości funkcji $W(k)$ wynika, że dla każdego stanu o wektorze falowym k, istnieje stan symetryczny o wektorze $-k$. Wypływa stąd wniosek, że wektor falowy dziury jest co do wartości równy wektorowi falowemu wzbudzonego elektronu, ma natomiast w stosunku do niego przeciwny kierunek

$$k_d = -k_e \tag{6/48}$$

Widać to bezpośrednio na rys. 6/20. Ze względu na to, że przyczynki wnoszone do sumy (6/46) przez elektrony znajdujące się w symetrycznie położonych stanach znoszą się wzajemnie, wypadkowy pęd wszystkich elektronów w pasmie walencyjnym jest równy pędowi jedynego występującego w tym pasmie „elektronu bez pary".

Podobnie energia dziury jest równa energii wszystkich N-1 elektronów w pasmie walencyjnym. Analogicznie do wzoru (6/47) możemy więc napisać

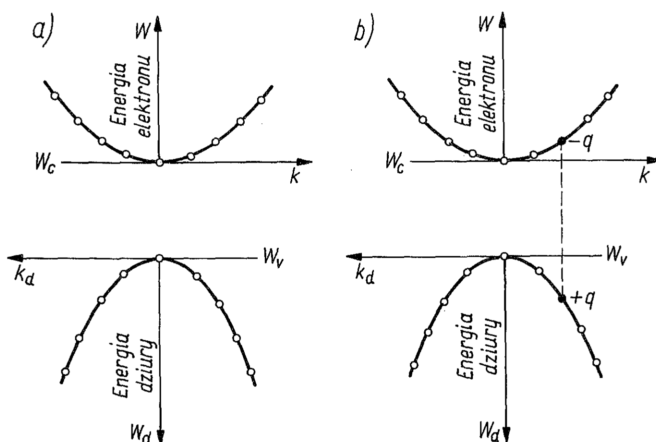
$$W_d(k) = \sum_{i=1}^{N} W_i - W_e(k) \tag{6/49}$$

[1] Masa efektywna elektronu jest niezależna od energii tylko w tej części pasma, dla której dopuszczalne jest przybliżenie paraboliczne.

Rozróżnianie między k_e i k_d jest w tym przypadku zbędne, ponieważ energia nie zależy od zwrotu wektora $\boldsymbol{k}$.

Pierwszy składnik w wyrażeniu (6/49) jest wielkością stałą, jest to bowiem energia wszystkich elektronów w całkowicie zapełnionym pasmie. Składnik ten można pominąć, ze względu na to, że poziom, względem którego jest mierzona energia, może być ustalony w sposób dowolny.

Ze wzoru (6/49) wynika, że im jest większa energia nie obsadzonego przez elektron stanu kwantowego w pasmie walencyjnym, tym jest mniejsza energia dziury. Wykres przedstawiający zależność dyspersyjną $W_d(k_d)$ dla dziur otrzymamy zatem z odpowiedniego wykresu dla elektronów w pasmie walencyjnym, zmieniając zwrot wektora falowego i kierunek osi energii – zob. rys. 6/21. Energia dziury jest tym większa, im niżej na rysunku znajduje się obsadzony przez nią stan kwantowy[1].



6/21 Zależności dyspersyjne dla elektronów i dla dziur: a) oba pasma puste; b) generacja (kreacja) jednej pary elektron-dziura

Z analizy tej wynika, że pasmo przewodnictwa i pasmo walencyjne można traktować symetrycznie, posługując się pojęciem dziury, która reprezentuje wypadkowe zachowanie się wszystkich N-1 elektronów w pasmie zawierającym N stanów kwantowych.

Całkowicie zapełnione pasmo walencyjne można więc traktować, operując językiem dziur, jako pasmo puste, tzn. nie zawierające dziur (rys. 6/21a), a wyrwanie elektronu z tego pasma – jako *kreację* dziury (rys. 6/21b).

Wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa nazywamy zatem *generacją* (kreacją) *pary elektron-dziura.*

Przy analizie zjawisk zachodzących w półprzewodnikach traktuje się dziury w pasmie walencyjnym w sposób analogiczny jak elektrony w pasmie prze-

[1] Podobnie energia piłki zanurzonej w wodzie staje się tym większa, im głębiej zostanie ona wciśnięta pod powierzchnię wody.

wodnictwa. Uważa się je więc za *nośniki ładunku*, których własności dynamiczne są reprezentowane przez masę efektywną.

Z kształtu funkcji $W_d(k_d)$ wynika, że masa efektywna dziury jest dodatnia i równa, co do wartości bezwzględnej, masie efektywnej elektronu w pasmie walencyjnym, tzn.

$$m_d^*(k) = -m_1^*(k) \tag{6/50}$$

Ładunek elektryczny dziury jest także dodatni, co wynika z następującego rozumowania: zakładając, że w rozważanym zakresie energii pasmo walencyjne jest paraboliczne, można skorzystać ze wzoru (6/45). Dla brakującego w tym pasmie elektronu otrzymamy wówczas

$$m_1^* v_e = \hbar k_e \tag{6/51}$$

dla dziury natomiast

$$m_d^* v_d = \hbar k_d \tag{6/52}$$

Opierając się na wzorach (6/48) i (6/50), dochodzimy do wniosku, że dziura porusza się z prędkością v_d równą (co do wartości i kierunku) prędkości v_e elektronu, który znajdowałby się w nie obsadzonym stanie w pasmie walencyjnym. Ponieważ prąd elektryczny wytwarzany przez elektrony całkowicie zapełnionego pasma jest równy zeru, zatem usunięcie z tego pasma 1 nośnika o ładunku $-q$ poruszającego się z prędkością v_e jest równoważne wprowadzeniu do tego pasma nośnika o ładunku $+q$ poruszającego się z tą samą prędkością[1].

Dla pasm parabolicznych oznaczamy masy efektywne elektronów w pasmie przewodnictwa i dziur w pasmie walencyjnym odpowiednio symbolami m_n^* i m_p^*. Obie te wielkości są dodatnie i niezależne od energii, a zastosowane indeksy wiążą się ze znakiem ładunku tych nośników: n – negatywny, p – pozytywny. Wprowadzając te oznaczenia, można wzór (6/43) przepisać w postaci

$$\boxed{W(k) = W_c + \frac{\hbar^2 k^2}{2m_n^*}} \tag{6/53}$$

Pierwszy składnik reprezentuje energię potencjalną, drugi zaś energię kinetyczną elektronu w pasmie przewodnictwa.

Analogiczny wzór dla pasma walencyjnego ma postać

$$\boxed{W(k) = W_v - \frac{\hbar^2 k^2}{2m_p^*}} \tag{6/54}$$

gdzie W – oznacza energię elektronu. Energię dziury wyrażamy natomiast wzorem

$$W_d(k) = W_v + \frac{\hbar^2 k^2}{2m_p^*} \tag{6/55}$$

[1] W rozumowaniu tym założyliśmy dla prostoty, że rozpatrywane pasmo jest paraboliczne. Nie ma to jednak istotnego znaczenia, a otrzymany wniosek dotyczący ładunku dziury jest ogólnie słuszny.

Tu również pierwszy składnik reprezentuje energię potencjalną, drugi zaś energię kinetyczną.

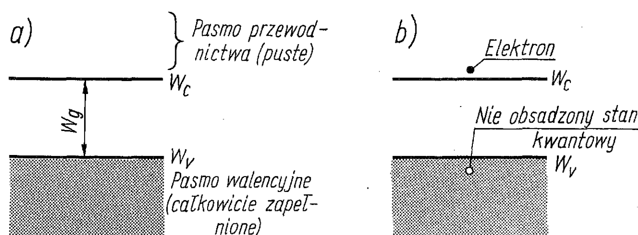
Dzięki wprowadzeniu pojęcia dziury pozbywamy się kłopotu z „dziwnymi" elektronami o ujemnej masie efektywnej. Dziura zachowuje się bowiem jak „normalna" cząstka, ma bowiem dodatnią masę efektywną. W sytuacji przedstawionej na rys. 6/21b dąży ona do tego, aby podobnie jak elektron w pasmie przewodnictwa, zająć stan kwantowy o najmniejszej energii. Proces ten opiszemy, operując najpierw językiem elektronów walencyjnych, a następnie językiem dziur.

Rozpatrując zachowanie się elektronów walencyjnych, posłużymy się rys. 6/20. W wyniku zderzeń tych elektronów z niedoskonałościami sieci krystalicznej będą one miały tendencję do „przeskakiwania" do stanów kwantowych o mniejszych energiach. Tak więc stan nieobsadzony zostanie zajęty przez elektron, znajdujący się w stanie położonym wyżej na rys. 6/20. Przechodząc do stanu o mniejszej energii elektron ten zwiększa swą prędkość. Widać to na rysunku, a wynika to z ujemnej masy efektywnej rozpatrywanych tu elektronów. Na opróżnione miejsce „przeskakuje" z kolei następny elektron z wyżej położonego stanu kwantowego. Proces ten trwa tak długo, aż w pasmie walencyjnym zostaną obsadzone wszystkie stany kwantowe z wyjątkiem położonego w punkcie $k = 0$ stanu o największej energii.

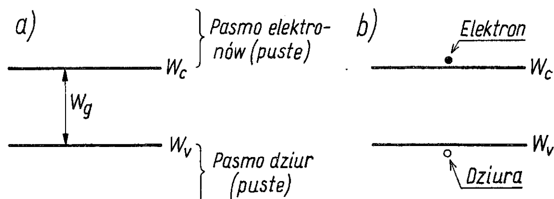
O ileż prościej przedstawia się opis tego procesu w języku dziur – rys. 6/21b. Dziura, podobnie jak elektron w pasmie przewodnictwa, zderza się z niedoskonałościami sieci krystalicznej. W wyniku tego traci ona stopniowo swą energię kinetyczną, czemu towarzyszy zmniejszanie się jej prędkości. W końcu dziura zatrzymuje się w stanie o najmniejszej energii, tzn. w stanie o $k = 0$.

Przy rozpatrywaniu działania większości elementów półprzewodnikowych nie zachodzi konieczność posługiwania się zależnościami dyspersyjnymi przedstawionymi na rys. 6/21. Do tego celu wystarcza najczęściej uproszczony model pasmowy półprzewodnika podany przykładowo na rys. 6/5. Na takim modelu można również pokazać wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa. Rozpatrując pasmo walencyjne można przy tym operować albo językiem elektronów (rys. 6/22), albo też językiem dziur (rys. 6/23).

W pierwszym przypadku mówimy, że przed wzbudzeniem pasmo walencyjne było całkowicie zapełnione przez elektrony, a po wzbudzeniu pojawia się w tym pasmie nieobsadzony stan kwantowy, w drugim zaś, że przed wzbudzeniem



6/22 Wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa: a) stan obsadzenia pasm przed wzbudzeniem; b) stan obsadzenia po wzbudzeniu



6/23 Generacja pary elektron-dziura: a) stan obsadzenia pasm przed generacją; b) stan obsadzenia po generacji

pasmo walencyjne było puste, a w wyniku wzbudzenia pojawia się w nim dziura — nośnik o dodatnim ładunku elektrycznym. W pewnych okolicznościach jeden sposób opisu jest dogodniejszy od drugiego, w innych zaś zachodzi sytuacja odwrotna.

Pojęcie dziury jest jedną z trudniejszych konstrukcji, jakimi posługuje się fizyka ciała stałego. Z tego względu pojęciu temu poświęciliśmy dość dużo miejsca. Na zakończenie warto jeszcze zwrócić uwagę na następującą sprawę: W literaturze operuje się często modelem dziury, w którym jest ona traktowana jako zerwane wiązanie przy określonym atomie sieci krystalicznej. Wskutek wytworzenia się „wolnego miejsca" staje się możliwy ruch elektronów tworzących wiązania międzyatomowe w krysztale. Ruch dziury jest w tym modelu traktowany jako „przeskakiwanie" zerwanego wiązania od atomu do atomu w wyniku przemieszczania się elektronów walencyjnych w przeciwnym kierunku.

Model ten jest bardzo poglądowy, zawiera jednakże pewną nieścisłość, przypisuje bowiem dziurze określoną lokalizację. Tymczasem dziurze o określonej wartości wektora falowego, zgodnie z zasadą nieoznaczoności Heisenberga, nie można przypisać żadnego wyróżnionego miejsca w krysztale. Odnosi się to również do elektronu o określonej wartości pędu. W punkcie 6.2 przyjmowaliśmy, że prawdopodobieństwo znalezienia takiego elektronu jest w każdej komórce elementarnej jednakowe. Na tej podstawie nadaliśmy funkcji falowej elektronu postać funkcji Blocha, co doprowadziło nas do modelu pasmowego. Konsekwencją tego modelu jest delokalizacja elektronów i dziur w półprzewodniku. „Przeskakiwanie" na wolne miejsca zachodzi, jak o tym już mówiliśmy, w przestrzeni wektora falowego, a nie w przestrzeni rzeczywistej. Jest to zasadnicza różnica, z której warto jest zdawać sobie sprawę.

7

PÓŁPRZEWODNIKI

W rozdziale tym zapoznamy się ze strukturą energetyczną półprzewodników ze szczególnym uwzględnieniem tych materiałów, które znajdują lub też w niedalekiej przyszłości mogą znaleźć szerokie zastosowanie w elementach półprzewodnikowych. Struktura krystaliczna realnych półprzewodników jest znacznie bardziej skomplikowana od jednowymiarowego modelu, którym posługiwaliśmy się w poprzednim rozdziale. Mimo to, rezultaty uzyskane dla modelu jednowymiarowego dadzą się w pewnym stopniu uogólnić na przypadek trójwymiarowy i umożliwiają łatwiejsze zrozumienie struktury pasmowej rzeczywistych kryształów.

7.1 PODSTAWOWE POJĘCIA I ZALEŻNOŚCI

Z rozważań zawartych w poprzednim rozdziale wynika, że wszystkie informacje dotyczące własności dynamicznych elektronu w ciele stałym są zawarte w jego zależności dyspersyjnej, tj. funkcji

$$W = W(\boldsymbol{k}) \tag{7/1}$$

opisującej zależność energii elektronu od wektora falowego. O funkcji tej mówimy, że przedstawia ona rozkład energii w *przestrzeni wektora falowego* (lub krótko: w *przestrzeni* $\boldsymbol{k}$), tzn. w przestrzeni, w której współrzędnymi są składowe wektora $\boldsymbol{k}$. Podobnie jak w rozpatrywanym w p. 6.2 modelu jednowymiarowym, kierunki osi współrzędnych w przestrzeni k są prostopadłe do odpowiednich płaszczyzn sieciowych kryształu. W ogólnym przypadku tworzą one zatem ukośnokątny układ współrzędnych.

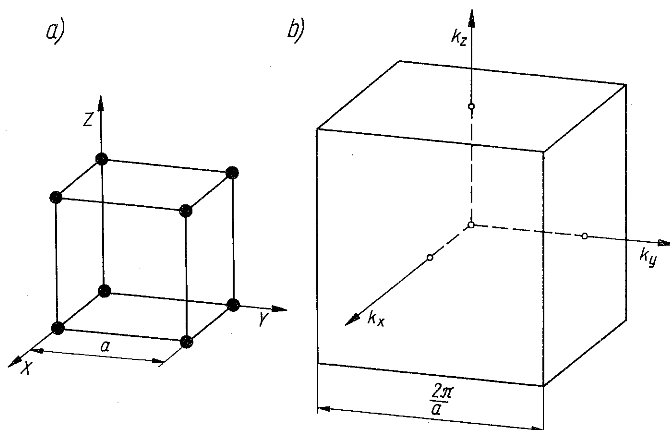
Olbrzymia większość interesujących nas półprzewodników krystalizuje jednakże w układzie regularnym. Osie współrzędnych w przestrzeni k są wówczas równoległe do osi krystalograficznych X, Y, Z, tworzą więc układ prostokątny –

zob. rys. 7/1. Zależność (7/1) można wówczas przedstawić w postaci

$$W = W(k_x, k_y, k_z) \tag{7/2}$$

Funkcja ta ma ogólne własności analogiczne do odpowiedniej funkcji jednowymiarowej omawianej w p. 6.4. Spośród tych własności na przypomnienie zasługuje przede wszystkim jej parzystość oraz periodyczność. Okres identyczności jest jednakże różny dla różnych kierunków wektora $\boldsymbol{k}$. Dzięki periodyczności funkcji (7/2) możemy ograniczyć nasze rozważania do pewnego obszaru przestrzeni k, który pokrywa całkowicie zakres zmienności funkcji $W(\boldsymbol{k})$. Obszar ten nazywamy, podobnie jak w p. 6.4, pierwszą strefą Brillouina.

Kształt tej strefy zależy od symetrii rozważanej struktury krystalicznej. W najprostszym przypadku sieci regularnej prostej (niecentrowanej), której komórką elementarną jest sześcian o boku a (rys. 7/1a), strefa Brillouina jest sześcianem o boku $2\pi/a$ (rys. 7/1b). Jest to zwykłe uogólnienie rozpatrywanego w p. 6.4 przypadku jednowymiarowego. Podobnie jak tam, środek strefy Brillouina znajduje się w początku układu współrzędnych, tzn. w punkcie $\boldsymbol{k} = 0$.



7/1 Komórka elementarna (a) i pierwsza strefa Brillouina (b) dla sieci regularnej prostej

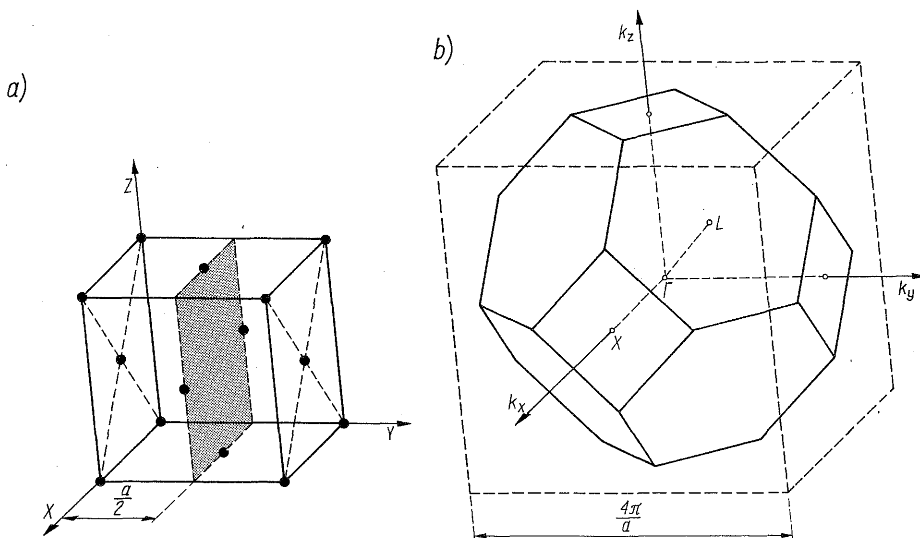
W fizyce półprzewodników największe znaczenie ma sieć regularna płasko centrowana (rys. 7/2a). Taką sieć mają dwie najpopularniejsze struktury krystaliczne: struktura diamentu i struktura blendy cynkowej. Dla tej sieci strefa Brillouina jest czternastościanem przedstawionym na rys. 7/2b. Bryłę tę można sobie wyobrazić jako otrzymaną z sześcianu o boku $4\pi/a$ przez odcięcie wszystkich naroży. Ma ona więc 6 ścian kwadratowych oraz 8 ścian o postaci sześciokątów foremnych.

W strefie tej wyróżnia się pewne kierunki o szczególnie wysokim stopniu symetrii oraz pewne charakterystyczne punkty, którym nadaje się umowne nazwy. Spośród nich największe znaczenie mają następujące punkty (i wyznaczone przez nie kierunki):

– *punkt* Γ, leżący w środku strefy Brillouina, czyli w punkcie $\boldsymbol{k} = 0$;

— *punkt X*, leżący w środku ściany kwadratowej. Jak wynika z rys. 7/2b, istnieje 6 równoważnych punktów X, leżących na osiach k_x, k_y, k_z w odległości równej $2\pi/a$ od punktu Γ. Wyznaczają one 6 równoważnych kierunków [100];

— *punkt L*, leżący w środku ściany sześciokątnej. Istnieje 8 równoważnych punktów L, leżących w odległości równej $\frac{\pi\sqrt{3}}{a}$ od punktu Γ. Wyznaczają one 8 równoważnych kierunków [111].

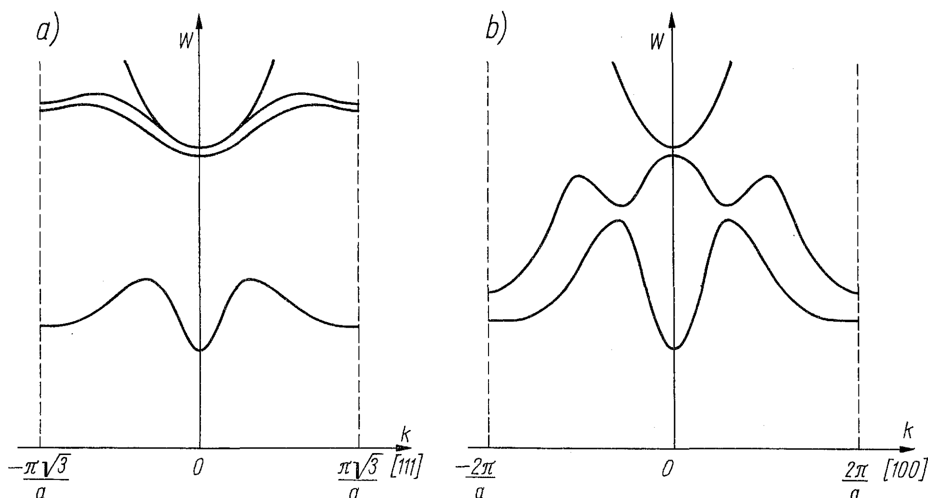


7/2 Komórka elementarna (a) i pierwsza strefa Brillouina (b) dla sieci regularnej płasko centrowanej

W celu zobrazowania struktury energetycznej półprzewodnika, przedstawia się ją w postaci zależności $W(\boldsymbol{k})$ dla wybranych kierunków wektora falowego. Przykładem takiego postępowania jest rys. 7/3, na którym pokazano strukturę energetyczną pasma przewodnictwa arsenku galu. Dla tego półprzewodnika, podobnie jak dla przeważającej większości kryształów, ekstrema funkcji $W(\boldsymbol{k})$ występują bądź w punkcie Γ, bądź też w osiach ΓX lub ΓL.

Jak wynika z rys. 7/3, dla kryształów 3-wymiarowych zależność $W(\boldsymbol{k})$ jest znacznie bardziej skomplikowana niż otrzymana dla 1-wymiarowego modelu Kroniga-Penney'a. Obok ekstremów w środku i na brzegach strefy Brillouina występują tu dodatkowe ekstrema w punktach pośrednich.

Ze względu na parzystość funkcji $W(\boldsymbol{k})$ do zobrazowania struktury energetycznej półprzewodnika wystarcza podanie połówek wykresów przedstawionych na rys. 7/3. W literaturze przyjęto, że strukturę pasmową półprzewodnika pokazuje się na wykresie, który powstaje z połączenia lewej połówki rys. 7/3a i prawej połówki rys. 7/3b. Wykres taki dla GaAs przedstawiono na rys. 7/6.



7/3 Pasmo przewodnictwa arsenku galu: a) przekrój strefy Brillouina w kierunku [111], b) przekrój w kierunku [100]

Inny sposób przedstawienia struktury energetycznej półprzewodnika polega na pokazaniu powierzchni izoenergetycznych w przestrzeni wektora $\boldsymbol{k}$. Dla elektronu swobodnego dokonano tego na rys. 6/8. Dla elektronu w krysztale są to z reguły dość skomplikowane powierzchnie zamknięte. Często posługujemy się przekrojami tych powierzchni płaszczyzną przechodzącą przez punkt Γ. Dla sieci regularnej płasko centrowanej rzutem stref Brillouina na płaszczyznę przesuniętą przez osie k_x i k_y jest ośmiokąt foremny. Krzywe izoenergetyczne w tej płaszczyźnie rysuje się często na tle tego ośmiokąta – zob. rys. 7/7.

Jak stwierdziliśmy w rozdz. 6, ważnym parametrem dynamicznym elektronu w krysztale jest jego masa efektywna. Wiąże ona przyspieszenie elektronu z działającą na niego siłą zewnętrzną. W modelu 1-wymiarowym zależność ta była wyrażona równaniem (6/39). Dla kryształu 3-wymiarowego związek między przyspieszeniem a siłą zewnętrzną jest na ogół znacznie będziej skomplikowany, bowiem każda ze składowych wektora ($\mathrm{d}v_x/\mathrm{d}t$, $\mathrm{d}v_y/\mathrm{d}t$, $\mathrm{d}v_z/\mathrm{d}t$) jest liniowo zależna od wszystkich 3 składowych wektora siły. W zapisie macierzowym zależność tę można wyrazić w następujący sposób:

$$\begin{bmatrix} \dfrac{\mathrm{d}v_x}{\mathrm{d}t} \\[2ex] \dfrac{\mathrm{d}v_y}{\mathrm{d}t} \\[2ex] \dfrac{\mathrm{d}v_z}{\mathrm{d}t} \end{bmatrix} = \frac{1}{\hbar^2} \begin{bmatrix} \dfrac{\partial^2 W}{\partial k_x^2} & \dfrac{\partial^2 W}{\partial k_x \partial k_y} & \dfrac{\partial^2 W}{\partial k_x \partial k_z} \\[2ex] \dfrac{\partial^2 W}{\partial k_y \partial k_x} & \dfrac{\partial^2 W}{\partial k_y^2} & \dfrac{\partial^2 W}{\partial k_y \partial k_z} \\[2ex] \dfrac{\partial^2 W}{\partial k_z \partial k_x} & \dfrac{\partial^2 W}{\partial k_z \partial k_y} & \dfrac{\partial^2 W}{\partial k_z^2} \end{bmatrix} \cdot \begin{bmatrix} F_x \\[2ex] F_y \\[2ex] F_z \end{bmatrix} \qquad (7/3)$$

Występująca tu macierz, składająca się z 9 pochodnych cząstkowych drugiego rzędu po składowych wektora falowego, podzielona przez $\hbar^2$ nosi nazwę *tensora odwrotności masy efektywnej*[1].

Z postaci tej zależności wynika, że w ogólnym przypadku kierunek wektora przyspieszenia nie pokrywa się z kierunkiem wektora siły. Dla sieci regularnej można przez odpowiedni dobór osi współrzędnych sprowadzić tensor odwrotności masy efektywnej do tzw. *postaci diagonalnej*, w której wszystkie pochodne mieszane są równe zeru. Oznaczając ten tensor symbolem $1/m^*$, można wówczas napisać

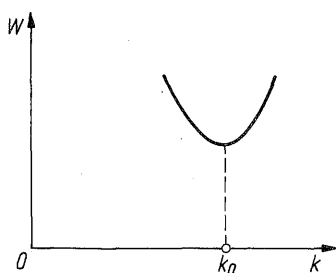
$$\frac{1}{m^*} = \begin{bmatrix} 1/m_1 & 0 & 0 \\ 0 & 1/m_2 & 0 \\ 0 & 0 & 1/m_3 \end{bmatrix} \tag{7/4}$$

Zamiast 9 składowych mamy tu tylko 3 różne od zera składowe. W niektórych kryształach występują pasma energetyczne, dla których te 3 składowe są sobie równe. Masa efektywna elektronu jest wówczas *skalarem*, a energia elektronu zależy tylko od wartości, nie zależy zaś od kierunku wektora falowego. Funkcja $W(\boldsymbol{k})$ wykazuje wówczas symetrię sferyczną, dlatego też pasma, dla których $m_1 = m_2 = m_3$ nazywamy *pasmami sferycznymi*.

Jeżeli w danym pasmie energetycznym występuje ekstremum w punkcie $\boldsymbol{k}_0$ (rys. 7/4), to w pobliżu tego ekstremum, podobnie jak w p. 6.6, można funkcję $W(\boldsymbol{k})$ rozwinąć w szereg potęgowy zawierający jedynie pochodne rzędu parzystego. W wielu przypadkach można wówczas dla pewnego odcinka tego pasma założyć jego paraboliczność. Funkcja $W(\boldsymbol{k})$ może być wówczas przedstawiona w postaci

$$W(\boldsymbol{k}) = W(\boldsymbol{k}_0) + \frac{\hbar^2}{2}\left(\frac{\varkappa_1^2}{m_1} + \frac{\varkappa_2^2}{m^2} + \frac{\varkappa_3^2}{m_3}\right) \tag{7/5}$$

gdzie $\boldsymbol{\varkappa} = \boldsymbol{k} - \boldsymbol{k}_0$. Wielkości $\varkappa_1, \varkappa_2, \varkappa_3$ są to składowe wektora $\boldsymbol{\varkappa}$ w układzie współrzędnych, w którym tensor odwrotności masy efektywnej wyraża się w postaci (7/4).



7/4
Fragment struktury energetycznej półprzewodnika, dla którego minimum pasma przewodnictwa występuje w punkcie $k_0 \neq 0$

[1] *Tensor* jest to pojęcie matematyczne, będące uogólnieniem skalara i wektora. Skalar, jak wiadomo, jest to wielkość charakteryzująca się jedynie wartością liczbową, nie zmieniająca się przy zmianie układu współrzędnych. Wektor jest to wielkość, którą można odwzorować za pomocą odcinka skierowanego. Ma on zatem określoną wartość, kierunek i zwrot. Tensor natomiast jest to operator liniowy, który działając na wektor daje w wyniku inny wektor. Może on być przedstawiony w postaci macierzy o 3 wierszach i 3 kolumnach. Jest to tzw. tensor drugiego rzędu. Wektory można uważać za tensory pierwszego rzędu, skalary zaś — za tensory rzędu zerowego.

Przybliżenie takie jest słuszne tylko dla małych wartości $\varkappa$. Przy większych odchyleniach od punktu k_0 zaczynają odgrywać rolę tzw. *efekty nieparaboliczności*. Będzie o nich mowa w p. 7.3.

Jak wynika ze wzoru (7/5), dla pasma parabolicznego powierzchnie izoenergetyczne są elipsoidami – w ogólnym przypadku trójosiowymi. Często zdarza się jednak, że $m_1 = m_2$, wówczas stają się one elipsoidami obrotowymi, a w szczególnym przypadku, gdy $m_1 = m_2 = m_3$ – sferami. Pasmo jest wówczas nie tylko paraboliczne, lecz również sferyczne.

Jeżeli rozpatrywane ekstremum funkcji $W(k)$ nie leży w punkcie Γ (tzn. jeżeli $k_0 \neq 0$), to struktura energetyczna zawiera w obrębie strefy Brillouina kilka równoważnych ekstremów występujących na równoważnych kierunkach przestrzeni k.

7.2 RODZAJE PÓŁPRZEWODNIKÓW

Półprzewodniki można sklasyfikować w różny sposób w zależności od przyjętego wyróżnika klasyfikacyjnego. Ze względu na skład chemiczny dzieli się je na:
- półprzewodniki pierwiastkowe,
- związki chemiczne,
- kryształy mieszane.

Półprzewodniki *pierwiastkowe* są zbudowane z atomów jednego pierwiastka. Zaliczamy do nich omawiane w p. 6.1.2 pierwiastki należące do IV grupy układu okresowego, a ponadto bor, selen i tellur – zob. rys. 7/5.

III	IV	V	VI	VII
B	C	N	O	F
Al	Si	P	S	Cl
Ga	Ge	As	Se	Br
In	Sn	Sb	Te	J
Tl	Pb	Bi	Po	At

7/5
Fragment układu okresowego pierwiastków z zaznaczonymi półprzewodnikami (pole zacienione) i metalami (pole zakreskowane)

Półprzewodnikowe *związki chemiczne* są najczęściej zbudowane z atomów 2 pierwiastków, przy czym ich skład ilościowy jest zgodny z wymaganiami wartościowości (jest to tzw. *skład stechiometryczny*). Przykładem takiego półprzewodnika jest rozpatrywany w p. 5.5 arsenek galu GaAs. Sieć krystaliczna tego półprzewodnika składa się z 2 wsuniętych w siebie regularnych podsieci, z których jedna jest utworzona z atomów galu, druga zaś z atomów arsenu. W wyniku tego nie tylko liczby atomów obu pierwiastków są dokładnie takie same, ale także rozkład atomów Ga i As jest doskonale regularny, kryształ zaś – idealnie jednorodny.

W zależności od przynależności pierwiastków tworzących związek chemiczny do grup układu okresowego, rozróżnia się związki typu:

IV–IV – do których należy np. SiC,
III–V – np. GaAs, InSb,
II–VI – np. ZnS, CdSe, HgTe,
IV–VI – np. PbS.

i inne (np. Cu_2O). Ważniejsze rodzaje półprzewodników tego typu zestawiono w dodatku D. Podano tam również dla każdego z nich wartość przerwy energetycznej.

Kryształy mieszane są zazwyczaj zbudowane z atomów 2 pierwiastków, albo też 2 lub więcej związków chemicznych. Ich skład ilościowy może zmieniać się w szerokich granicach. Przykładem takiego półprzewodnika jest kryształ mieszany $Ge_x\,Si_{1-x}$. Użyty tu wskaźnik x, który może przyjmować dowolną wartość w granicach od 0 do 1, jest równy stosunkowi liczby atomów germanu do liczby wszystkich atomów kryształu. Kryształy mieszane nie są idealnie jednorodne, ponieważ rozkład tworzących je atomów jest w pewnym stopniu przypadkowy.

Obok kryształów dwuskładnikowych (takich np. jak $Ge_x\,Si_{1-x}$) ważną rolę w elektronice półprzewodnikowej odgrywają kryształy trójskładnikowe i czteroskładnikowe. Przykładem kryształu trójskładnikowego może być arseno-fosforek galu $GaAs_{1-x}\,P_x$. Kryształ ten jest mieszaniną związków GaAs i GaP, co można również wyrazić za pomocą symbolu $(GaAs)_{1-x}(GaP)_x$. Półprzewodnik ten jest powszechnie stosowany do wyrobu diod elektroluminescencyjnych.

Jako przykład kryształu czteroskładnikowego można podać $Ga_{0,13}\,In_{0,87}\,As_{0,37}\,P_{0,63}$ stosowany np. w tzw. diodach Gunna. Kryształ ten jest mieszaniną 4 związków chemicznych: GaAs, GaP, InAs i InP.

Technologia nowych materiałów półprzewodnikowych, a w szczególności kryształów mieszanych, ma duże perspektywy rozwoju. W ostatnich latach pojawiają się wciąż nowe rodzaje kryształów o doskonalszych parametrach, bądź też dotychczas niespotykanych właściwościach.

Ze względu na własności fizyczne dzieli się półprzewodniki na samoistne i domieszkowe. W półprzewodnikach *samoistnych* w warunkach równowagi termodynamicznej[1] elektrony w pasmie przewodnictwa pojawiają się wyłącznie (a ściślej – prawie wyłącznie) wskutek wzbudzenia z pasma walencyjnego. Wzbudzeniu takiemu towarzyszy, jak wiadomo, pojawienie się dziury w pasmie walencyjnym. Skoro zatem wszystkie elektrony w pasmie przewodnictwa pochodzą z pasma walencyjnego, w półprzewodniku samoistnym koncentracje równowagowe elektronów (n_0) i dziur (p_0) są jednakowe

$$\boxed{n_0 = p_0} \qquad (7/6)$$

Idealny półprzewodnik samoistny charakteryzuje się doskonałą symetrią sieci krystalicznej, nie zawiera więc obcych atomów (zanieczyszczeń i domieszek), ani też innych defektów strukturalnych.

---

[1] Stan równowagi termodynamicznej charakteryzuje się tym, że nie występują żadne oddziaływania między półprzewodnikiem a otoczeniem (zob. p. 8.1).

Półprzewodniki *domieszkowe* zawierają celowo wprowadzone obce atomy (domieszki), które mają istotny wpływ na koncentracje równowagowe nośników ładunku w pasmie przewodnictwa i w pasmie walencyjnym.

Jak zobaczymy w rozdz. 8, w wyniku wprowadzenia domieszek może nastąpić zarówno zwiększenie, jak i zmniejszenie koncentracji elektronów w pasmie przewodnictwa. Zwiększeniu koncentracji tych elektronów towarzyszy przy tym zmniejszenie koncentracji dziur w pasmie walencyjnym i odwrotnie. W półprzewodniku domieszkowym dominuje więc jeden rodzaj nośników. Jeżeli są nimi elektrony, to półprzewodnik taki nazywamy *półprzewodnikiem o przewodnictwie elektronowym*, lub *półprzewodnikiem typu n*. W przeciwnym razie mówimy *o półprzewodniku o przewodnictwie dziurowym*, czyli o *półprzewodniku typu p*.

Posługując się pojęciem „domieszki" w definicji półprzewodnika domieszkowego, należy je pojmować bardzo szeroko. Rolę „domieszki" w tym sensie mogą spełniać różne defekty strukturalne, np. luki węzłowe, atomy własne w położeniach międzywęzłowych, a w niektórych przypadkach nawet atomy własne w położeniach węzłowych, np. w arsenku galu atom galu w położeniu, które powinno być obsadzone przez atom arsenu, lub odwrotnie[1].

Przejście od półprzewodnika domieszkowego do samoistnego zachodzi w sposób płynny. Tak więc ten sam kryształ może w niskich temperaturach wykazywać własności półprzewodnika domieszkowego, w wysokich natomiast — półprzewodnika samoistnego. Zagadnienia te będą rozpatrywane dokładniej w następnym rozdziale.

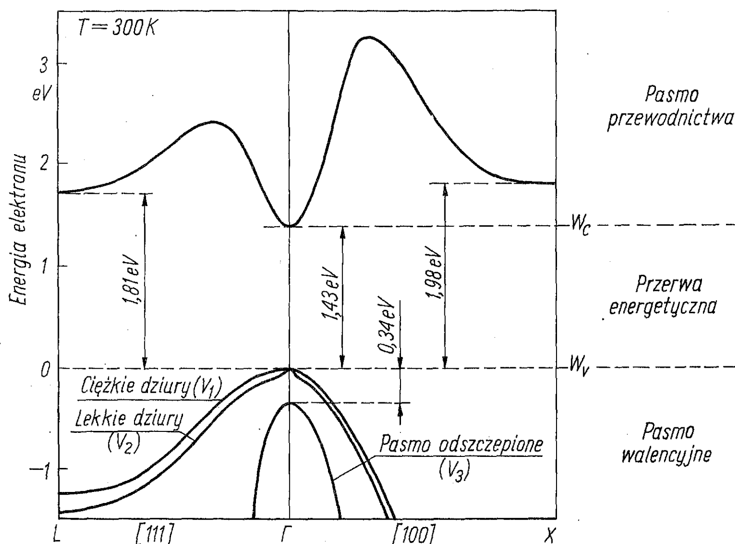
Ostatnim rodzajem klasyfikacji półprzewodników, z którym zapoznamy się w tej książce, jest podział na półprzewodniki z *prostą* i ze *skośną* przerwą energetyczną. Pojęcia te będą wyjaśnione w dalszych punktach tego rozdziału.

7.3 PÓŁPRZEWODNIKI Z PROSTĄ PRZERWĄ ENERGETYCZNĄ

W półprzewodnikach z prostą przerwą energetyczną najniżej położone minimum w pasmie przewodnictwa przypada w tym samym punkcie strefy Brillouina (tzn. przy tej samej wartości wektora $\boldsymbol{k}$), w którym występuje najwyżej położone maksimum w pasmie walencyjnym. Najczęściej (ale nie zawsze) jest to punkt Γ, w którym $\boldsymbol{k} = 0$. Typowym przykładem takiego półprzewodnika jest arsenek galu. Odgrywa on we współczesnej elektronice półprzewodnikowej bardzo ważną rolę, dlatego też zapoznamy się dokładnie z jego strukturą pasmową. Najważniejszy fragment tej struktury pokazano na rys. 7/6. Stanowi on, jak to już wyżej zaznaczono, złożenie 2 przekrojów strefy Brillouina wzdłuż kierunków ΓL i ΓX.

Z rysunku tego można odczytać, że najniższe minimum w pasmie przewodnictwa (dno tego pasma) występuje w punkcie Γ. W okolicy tego punktu pasmo przewodnictwa jest sferyczne i paraboliczne. Oznacza to, że masa efektywna elek-

[1] Tego rodzaju defekty strukturalne są w literaturze określane jako *„antysite defects"*. Termin „antysite" utworzony ze słów „anty" i angielskiego „site" (miejsce) nie ma odpowiednika w języku polskim.



7/6 Struktura pasmowa arsenku galu

tronów jest skalarna i nie zależy od energii elektronu. Wynosi ona

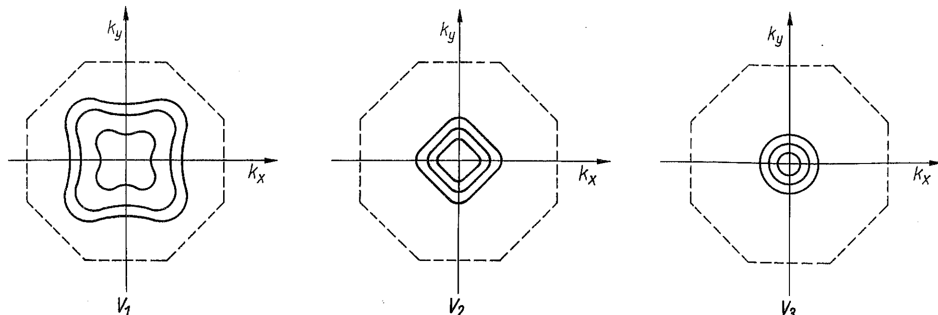
$$m_n^* = 0{,}067\ m_0$$

Przybliżenie paraboliczne jest słuszne w dość szerokim zakresie energii – do poziomu znajdującego się w odległości ok. 150 meV ponad dnem pasma przewodnictwa.

Poza minimum w punkcie Γ występują w tym pasmie dodatkowe minima w punktach L i X. Znajdują się one odpowiednio o 0,38 i 0,55 eV powyżej dna pasma przewodnictwa. Minima te nazywane są *dolinami bocznymi* w odróżnieniu od minimum w punkcie Γ, które nosi nazwę *doliny centralnej.* Doliny boczne mają znacznie mniejsze krzywizny, co świadczy o znacznie większych wartościach masy efektywnej elektronu w tych obszarach strefy Brillouina. Fakt ten ma istotny wpływ na zachowanie się elektronów poddanych działaniu silnego pola elektrycznego w arsenku galu i innych półprzewodnikach o podobnej strukturze pasma przewodnictwa (zob. p. 10.6.2).

Przechodząc z kolei do pasma walencyjnego, należy przede wszystkim zauważyć, że składa się ono z 3 niezależnych pasm, które można oznaczyć symbolami V_1, V_2 i V_3 – zob. rys. 7/6. Pasma V_1 i V_2 mają wspólne maksimum w punkcie Γ. Wyznacza ono wierzchołek pasma walencyjnego. Ze względu na różnicę krzywizny obu pasm, wyższe z nich nazywamy pasmem *ciężkich dziur,* niższe zaś – pasmem *lekkich dziur.*

Pasmo V_3 nosi nazwę *pasma odszczepionego.* Jego maksimum znajduje się również w punkcie Γ, jest jednakże przesunięte w dół o pewną wartość, która



7/7 Przekroje powierzchni izoenergetycznych dla pasm walencyjnych w arsenku galu

dla GaAs wynosi 0,34 eV. Przesunięcie to, nazywane *odszczepieniem* pasma, jest spowodowane oddziaływaniem spinowo-orbitalnym[1].

Pasma V_1 i V_2 nie są sferyczne. Powierzchnie izoenergetyczne w przestrzeni k są bardziej podobne do sześcianów niż do kul – zob. rys. 7/7. Pasmo V_3 jest natomiast sferyczne, ma ono jednak zwykle drugorzędne znaczenie.

Z kolei pasma V_1 i V_3 są paraboliczne w dość szerokim zakresie energii, natomiast pasmo V_2 tylko w bezpośrednim otoczeniu punktu Γ.

Wartości mas efektywnych uśrednione po wszystkich kierunkach wynoszą dla poszczególnych pasm odpowiednio[2]:

$$m_1^* = 0{,}50 m_0$$

$$m_2^* = 0{,}068 m_0$$

$$m_3^* = 0{,}133 m_0$$

W półprzewodnikach z prostą przerwą energetyczną możliwe są bezpośrednie przejścia elektronów z pasma walencyjnego do pasma przewodnictwa w wyniku absorpcji promieniowania elektromagnetycznego o dostatecznie dużej energii fotonów.

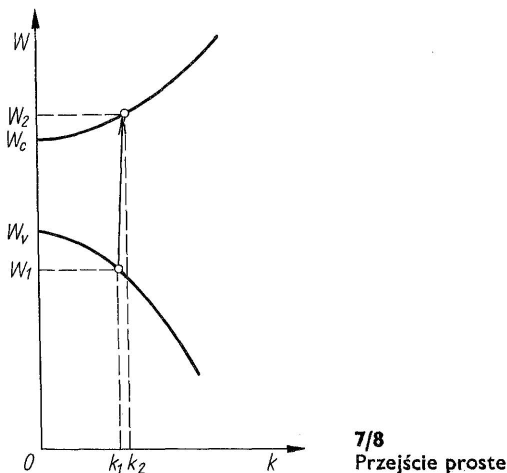
W przejściu takim musi być spełniona nie tylko zasada zachowania energii, lecz również zasada zachowania pędu. Oznaczając przez W_1 i k_1 wartość energii elektronu i jego wektora falowego w pasmie walencyjnym, a przez W_2 i k_2 te same wielkości w pasmie przewodnictwa (zob. rys. 7/8), można wymienione wyżej warunki zapisać następująco:

$$W_1 + h\nu = W_2 \qquad (7/7)$$

$$\hbar k_1 + \frac{h\nu}{c} = \hbar k_2 \qquad (7/8)$$

[1] *Oddziaływanie spinowo-orbitalne* powstaje w atomach i układach zbudowanych z atomów wskutek wzajemnego oddziaływania 2 momentów magnetycznych: momentu spinowego oraz momentu związanego z ruchem orbitalnym elektronu. Oddziaływanie to jest przyczyną występowania w widmach atomowych tzw. struktury subtelnej, o czym była mowa w p. 3.3.

[2] Zastosowane tu oznaczenia m_1^*, m_2^* i m_3^* nawiązują do odpowiednich oznaczeń pasm V_1, V_2 i V_3, nie mają natomiast nic wspólnego z masami efektywnymi występującymi w wyrażeniu (7/4).



7/8
Przejście proste

Z pierwszego z tych równań wynika, że energia fotonu musi być większa niż szerokość przerwy energetycznej półprzewodnika, a więc

$$\boxed{h\nu > W_g} \tag{7/9}$$

Z drugiego równania natomiast można wyciągnąć wniosek, że rozważane tu przejście odbywa się prawie bez zmiany wektora falowego. Łatwo bowiem można stwierdzić, że pęd fotonu spełniającego z niewielką nadwyżką warunek (7/9) jest pomijalnie mały w stosunku do pędu krystalicznego na granicy strefy Brillouina.

Problem 7.1. Wyznacz pęd fotonu o energii $h\nu = W_g$ dla arsenku galu i stwierdź, ile razy jest on mniejszy od pędu krystalicznego na granicy Brillouina strefy tego półprzewodnika.

Jak już mówiliśmy w p. 6.7, przejścia elektronowe, przy których nie zachodzi istotna zmiana wektora falowego, nazywamy *przejściami prostymi*. Dla promieniowania elektromagnetycznego spełniającego warunek (7/9) prawdopodobieństwo absorpcji fotonu, a co za tym idzie prawdopodobieństwo takich przejść w próbce półprzewodnika jest bardzo duże. Miarą intensywności przejść optycznych jest współczynnik absorpcji promieniowania zdefiniowany wzorem

$$\alpha(\nu) = -\frac{1}{S(x)}\frac{\mathrm{d}S}{\mathrm{d}x} \tag{7/10}$$

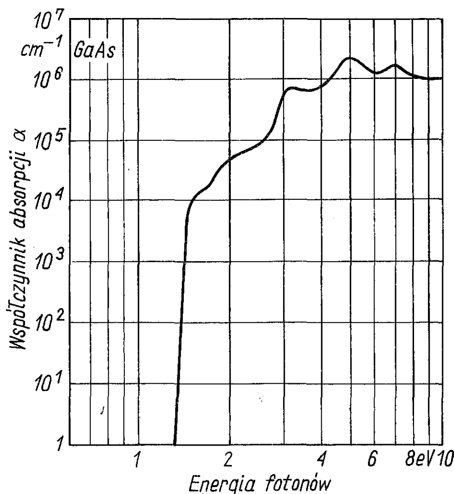
gdzie: $S(x)$ — natężenie fali elektromagnetycznej w punkcie x[1], x zaś — współrzędna mierzona w kierunku rozchodzenia się fali. Współczynnik absorpcji jest to więc względna szybkość zmniejszania się natężenia fali elektromagnetycznej wzdłuż drogi propagacji tej fali.

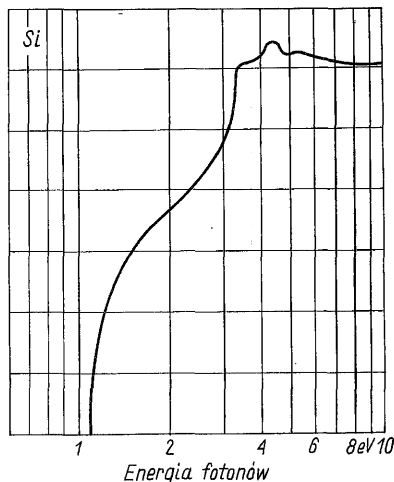
Współczynnik ten zależy od częstotliwości fali elektromagnetycznej. Przykład takiej zależności pokazano na rys. 7/9a. Dla fotonów o energii $h\nu \leqslant W_g$

---

[1] Wielkość ta była zdefiniowana w p. 1.2.

absorpcja prawie nie występuje – kryształ jest więc przezroczysty[1]. Po przekroczeniu punktu $h\nu = W_g$ następuje niezwykle szybki wzrost współczynnika absorpcji – jest to tzw. *krawędź absorpcji.* Współczynnik absorpcji osiąga bardzo szybko wartość ok. 10^4 cm^{-1}. Natężenie fali elektromagnetycznej maleje wówczas e-krotnie na drodze o długości 10^{-4} cm, czyli 1 µm. Na odcinku o długości 10 µm natężenie fali maleje wówczas ponad 20 000 razy, a zatem nawet niezwykle cienkie próbki półprzewodnika stają się całkowicie nieprzezroczyste dla promieniowania o takiej częstotliwości.





7/9 Zależność współczynnika absorpcji arsenku galu i krzemu od energii fotonów

Wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa jest procesem odwracalnym. Procesowi odwrotnemu, tzn. przejściu elektronu z pasma przewodnictwa do pasma walencyjnego, towarzyszy emisja fotonu o energii $h\nu$, którą można wyznaczyć ze wzoru (7/7). Zjawisko to, nazywane *rekombinacją promienistą,* jest wykorzystywane w diodach elektroluminescencyjnych i laserach półprzewodnikowych.

Przejściu elektronu z pasma przewodnictwa do pasma walencyjnego (tzn. zanikowi pary elektron-dziura) nie musi jednakże towarzyszyć emisja fotonu. Istnieje wiele innych możliwych mechanizmów rozpraszania energii wydzielanej przez rekombinującą parę nośników. Zjawiska te nazywamy ogólnie *rekombinacją niepromienistą.* Jednym z możliwych mechanizmów rekombinacyjnych jest *zjawisko Augera*[2], polegające na tym, że energia wydzielająca się w procesie rekombinacji zostaje przekazana innemu nośnikowi (elektronowi lub dziurze), który

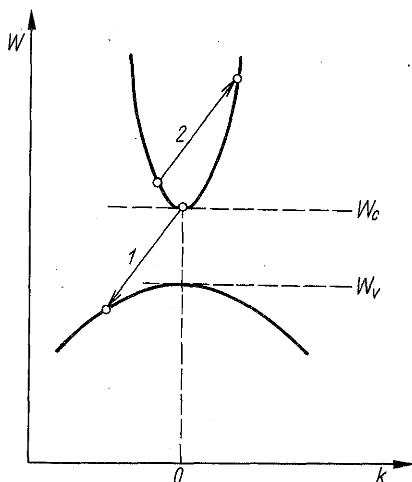
---

[1] Obserwuje się tu jedynie bardzo słabe tło absorpcyjne wywołane pochłanianiem fotonów, spowodowanym innymi szczątkowymi mechanizmami.

[2] Czytaj: ożé.

następnie rozprasza ją np. przez emisję pewnej liczby fononów. W procesie tym musi być również spełniona zasada zachowania energii i zasada zachowania pędu. Widzimy to na rys. 7/10, na którym pokazano jeden z możliwych procesów, które obejmujemy wspólną nazwą zjawisk Augera.

Omawiana wyżej struktura energetyczna arsenku galu jest typowa dla wielu związków typu III-V i II-VI o stosunkowo szerokiej przerwie energetycznej (rzędu 1 eV lub większej). Struktury pasmowe poszczególnych półprzewodników różnią się oczywiście wartością parametrów liczbowych, ale ogólny charakter zależności $W(\boldsymbol{k})$ jest bardzo podobny – zob. np. rys. 10/14.



7/10
Zjawisko Augera w półprzewodniku z prostą przerwą energetyczną
1 — elektron rekombinujący z dziurą;
2 — elektron wzbudzony na wyższy poziom energetyczny w pasmie przewodnictwa (elektron Augera)

Dla półprzewodników z wąską przerwą energetyczną (poniżej 0,5 eV) pojawia się problem *nieparaboliczności* pasm. Przyczyną tej nieparaboliczności jest silne oddziaływanie między pasmem przewodnictwa a energetycznie blisko położonym pasmem lekkich dziur.

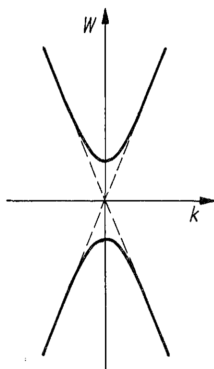
Dla takich półprzewodników jak np. InSb, InAs, HgTe i HgSe, w których ekstrema pasm występują w punkcie Γ, można pasma te uważać za sferyczne, a ich kształt z dobrym przybliżeniem opisać funkcją

$$W = \pm \sqrt{\left(\frac{W_g}{2}\right)^2 + \frac{\hbar^2 k^2}{2m^*} W_g} \qquad (7/11)$$

gdzie: W – energia mierzona względem środka przerwy energetycznej, m^* zaś – masa efektywna elektronów na dnie pasma przewodnictwa (równa masie efektywnej lekkich dziur na wierzchołku pasma walencyjnego). Przykładowo, dla InSb mamy $W_g = 0{,}18$ eV oraz $m^* = 0{,}013\, m_0$.

Zależność wyrażoną wzorem (7/11) przedstawiono na rys. 7/11. Przy małych wartościach wektora $\boldsymbol{k}$, gdy spełniony jest warunek

$$\frac{\hbar^2 k^2}{2m^*} \ll \frac{W_g}{4} \qquad (7/12)$$



7/11
Kształt pasm energetycznych w półprzewodniku z wąską przerwą energetyczną

można wyrażenie (7/11) sprowadzić do postaci

$$W \approx \pm \left(\frac{W_g}{2} + \frac{\hbar^2 k^2}{2m^*} \right) \tag{7/13}$$

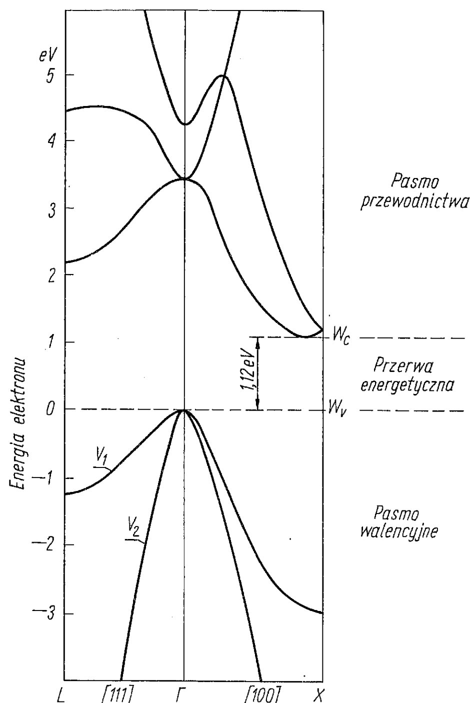
Wynika stąd, że zarówno pasmo przewodnictwa, jak i pasmo lekkich dziur, są w pobliżu swych krawędzi paraboliczne. W miarę wzrostu energii nośników ich masy efektywne wzrastają, przy czym w interesującym nas przedziale energii wzrost ten może być nawet kilkakrotny.

Innym przykładem półprzewodników z wąską przerwą energetyczną są związki ołowiu PbS, PbSe i PbTe, stosowane od kilkudziesięciu lat jako detektory promieniowania podczerwonego. W półprzewodnikach tych zarówno maksimum pasma walencyjnego, jak i minimum pasma przewodnictwa występują w punkcie L strefy Brillouina. Pasma te są nie tylko nieparaboliczne, ale także niesferyczne (powierzchnie izoenergetyczne są elipsoidami obrotowymi).

7.4 PÓŁPRZEWODNIKI ZE SKOŚNĄ PRZERWĄ ENERGETYCZNĄ

W wielu półprzewodnikach, takich np. jak krzem, german oraz fosforek galu (GaP) struktura pasma walencyjnego jest bardzo podobna do struktury tego pasma w omawianym już arsenku galu. Istotna różnica w stosunku do GaAs polega natomiast na tym, że minimum pasma przewodnictwa nie występuje w punkcie Γ, lecz w okolicy punktu X lub L strefy Brillouina. W krzemie (rys. 7/12) występuje ono na osi ΓX w punkcie o wartości $k = 0{,}85\ 2\pi/a$, w germanie — w punkcie L, a w GaP prawie dokładnie w punkcie X. O półprzewodnikach takich mówimy, że mają *skośną przerwę* energetyczną.

Występowanie minimum pasma przewodnictwa w innym punkcie strefy Brillouina niż maksimum pasma walencyjnego, ma istotny wpływ na własności optyczne i elektryczne półprzewodnika. Objawia się to przede wszystkim znacznie mniejszym prawdopodobieństwem przejść międzypasmowych w porównaniu z półprzewodnikami, w których ekstrema pasm występują przy tej samej wartości wektora falowego.

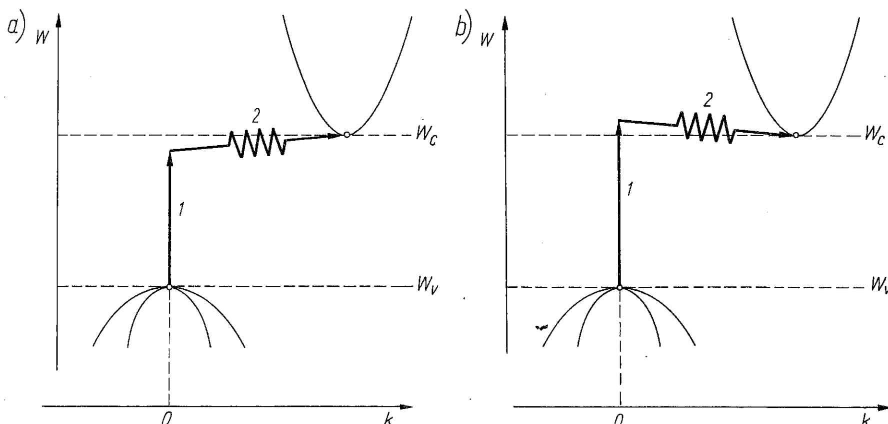


7/12 Struktura pasmowa krzemu (Opracowano na podstawie: J. R. Chelikowsky, M. L. Cohen: Electronic structure of silicon. *Phys. Rev.* **B**4 5095 (1974))

W półprzewodnikach ze skośną przerwą energetyczną bezpośrednie przejście elektronu między ekstremami pasm wiąże się, jak to widać na rys. 7/12, nie tylko ze zmianą energii elektronu, lecz również ze znaczną zmianą jego wektora falowego (czyli także jego pędu). Ze względu na to, że pęd fotonu jest znikomo mały w stosunku do wartości odpowiadającej promieniowi strefy Brillouina (tzn., w przypadku krzemu – w stosunku do wartości h/a), przejścia takie nie mogą zachodzić jedynie w wyniku absorpcji lub emisji fotonu.

Spełnienie zasady zachowania pędu staje się możliwe, gdy w procesie tym obok fotonu bierze udział dodatkowa cząstka, która dostarcza „brakującego" pędu. W półprzewodnikach cząstką tą jest z reguły fonon – kwant energii drgań sieci krystalicznej. Energie fononów są zwykle niewielkie (rzędu 10 meV) w stosunku do szerokości przerwy energetycznej (dla krzemu jest to np. wartość rzędu $1\% W_g$). Pęd fononu natomiast może być porównywalny z wartością odpowiadającą promieniowi strefy Brillouina.

Wzbudzenie elektronu z wierzchołka pasma walencyjnego na dno pasma przewodnictwa może więc nastąpić jedynie przy jednoczesnej absorpcji fotonu oraz absorpcji lub emisji fononu. Schematy tych przejść, zwanych *przejściami skośnymi*, pokazano na rys. 7/13[1]. Odpowiedniki warunków (7/7) i (7/8) można dla przejść skośnych sformułować następująco:



7/13 Poglądowy sposób przedstawienia przejść skośnych: a) przejście z absorpcją fononu; b) przejście z emisją fononu

$$\hbar\omega \pm \hbar\Omega \geqslant W_g \tag{7/14}$$

$$\pm\hbar K = \hbar\,\Delta k \tag{7/15}$$

gdzie: $\omega = 2\pi\nu$ – pulsacja promieniowania elektromagnetycznego (fotonu); Ω – pulsacja drgań sieci krystalicznej (fononu); K – wektor falowy fononu; $\hbar\,\Delta k$ – zmiana prądu elektronu. We wzorze (7/15) pominięto pęd fotonu jako znikowo mały w stosunku do pozostałych wielkości. Znak plus w tym wzorze dotyczy przypadku absorpcji, znak minus zaś – emisji fononu.

Przy przejściach odwrotnych (tzn. rekombinacji elektronu z dziurą) następuje jednoczesna emisja fotonu oraz emisja lub absorpcja fononu.

Zależność współczynnika absorpcji krzemu od energii fotonu pokazano na rys. 7/9b. W porównaniu z raptownym wzrostem tego współczynnika, który obserwowaliśmy dla arsenku galu po przekroczeniu energii równej W_g (rys. 9/7a), przebieg α ($h\nu$) dla krzemu jest znacznie łagodniejszy. Wynika to z faktu, że prawdopodobieństwo procesów trójcząstkowych (elektron+foton+fonon) jest wielokrotnie mniejsze niż prawdopodobieństwo procesu, w którym biorą udział tylko 2 cząstki (elektron i foton). Kształt krawędzi absorpcji krzemu zależy silnie od temperatury, gdyż zarówno liczba fononów, jak i rozkład ich energii są funkcjami temperatury.

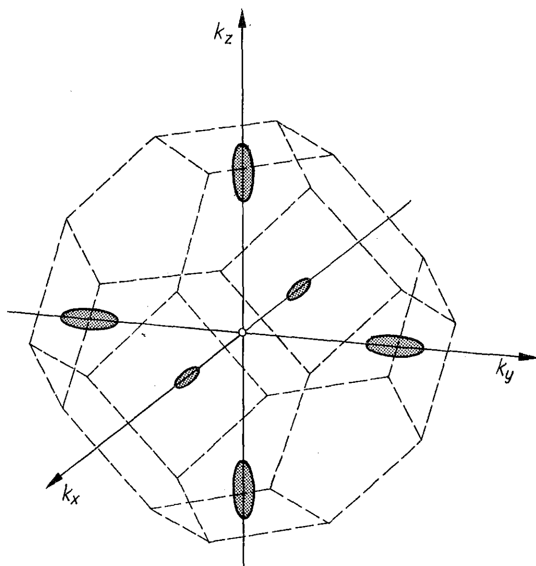
Drugą bardzo istotną konsekwencją występowania minimum pasma przewodnictwa poza punktem Γ jest niesferyczność tego pasma. Dla GaAs powierzchnia stałej energii w przestrzeni k jest powierzchnią kulistą, a masa efektywna elektronu w pasmie przewodnictwa — skalarem. W półprzewodnikach ze skośną przerwą energetyczną mamy do czynienia z sześcioma lub ośmioma równoważnymi minimami tego pasma rozłożonymi symetrycznie na

[1] W półprzewodniku ze skośną przerwą energetyczną mogą występować przejścia proste, wymagają one jednakże znacznie większych energii fotonów. Jak wynika z rys. 7/12, w krzemie przejścia takie stają się możliwe, gdy energia fotonu przekroczy wartość ok. 3,4 eV. Pojawienie się przejść prostych uwidacznia się na rys. 7/9b raptownym wzrostem współczynnika absorpcji przy tej wartości energii fotonów.

kierunkach [100] lub [111]. Powierzchnie izoenergetyczne mają w tych minimach kształt elipsoid obrotowych – zob. rys. 7/14. Masę efektywną elektronu wyznaczoną w kierunku osi symetrii elipsoidy nazywamy masą *podłużną* i oznaczamy symbolem $m^*_{n\|}$, a wyznaczoną w kierunku prostopadłym do osi symetrii – masą *poprzeczną* ($m^*_{n\perp}$). Dla krzemu masy te mają następujące wartości:

$$m^*_{n\|} = 0{,}916 m_0$$

$$m^*_{n\perp} = 0{,}191 m_0$$

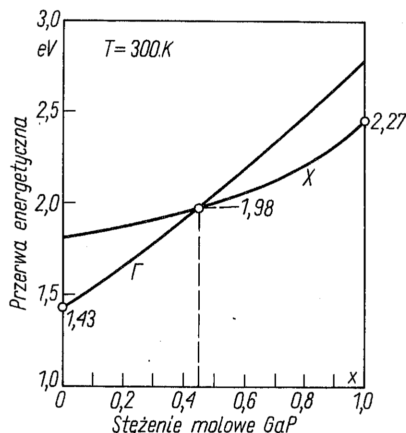


7/14
Powierzchnie izoenergetyczne krzemu

Interesującą możliwością technologiczną jest wytwarzanie kryształów mieszanych, w których, w zależności od składu, można w pewnych granicach „regulować" wartość przerwy energetycznej. Jako przykład takiego półprzewodnika rozważmy kryształ mieszany $GaAs_{1-x}P_x$. Jak pamiętamy, GaAs ma prostą przerwę energetyczną, przy czym minimum pasma przewodnictwa występuje w punkcie Γ, GaP natomiast ma przerwę skośną z minimum leżącym, praktycznie rzecz biorąc, w punkcie X.

W krysztale mieszanym obserwuje się prawie liniowe zmiany szerokości przerwy energetycznej w zależności od współczynnika x określającego skład kryształu – zob. rys. 7/15. Dla $x < 0{,}45$ przerwa energetyczna jest przerwą prostą i w zależności od wartości x szerokość jej zmienia się w granicach od 1,43 do 1,98 eV. Przy $x = 0{,}45$ minima w punktach Γ i X znajdują się na tej samej wysokości, a przy $x > 0{,}45$ mamy do czynienia z przerwą skośną, której szerokość zmienia się w granicach od 1,98 do 2,27 eV.

Z rozważań tych wynika, że stosując kryształy mieszane można wytwarzać „na zamówienie" półprzewodniki o żądanym rodzaju i żądanej szerokości przerwy energetycznej. Z możliwości tej korzysta się np. przy wytwarzaniu diod elektroluminescencyjnych. W zależności od składu kryształu $GaAs_{1-x}P_x$ otrzymuje się diody o barwie świecenia: czerwonej ($x = 0{,}4$), pomarańczowej ($x = 0{,}65$), żółtej ($x = 0{,}85$) lub zielonej ($x = 1$).

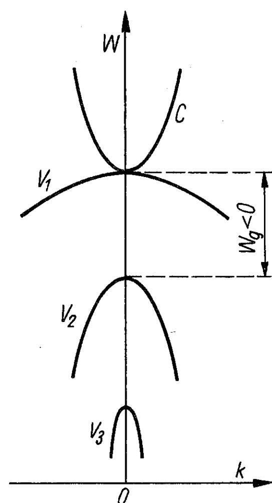


7/15
Przerwa energetyczna kryształu mieszanego $GaAs_{1-x}P_x$

7.5* PÓŁPRZEWODNIKI Z ZEROWĄ PRZERWĄ ENERGETYCZNĄ

Bardzo interesujące formy struktury pasmowej kryją się za nieco mylącymi nazwami „półprzewodniki z zerową (lub ujemną) przerwą energetyczną", jakie spotyka się w literaturze. W dotychczas rozważanych związkach półprzewodnikowych (takich np. jak GaAs) atomowe poziomy *s* kationu (Ga) znajdowały się powyżej atomowych poziomów *p* anionu (As). W wyniku tego te pierwsze były głównie odpowiedzialne za powstanie pasma przewodnictwa, drugie zaś – za powstanie pasma walencyjnego.

W kryształach zbudowanych z atomów o stosunkowo dużej liczbie atomowej, takich np. jak szara cyna lub HgTe, obserwuje się tzw. *odwrócenie* struktury pasmowej, czego przyczyną jest zachodzenie na siebie atomowych poziomów *s* i *p*. Przykład takiej *odwrotnej* struktury pasmowej przedstawiono na rys. 7/16.



7/16
Struktura pasmowa półprzewodnika z ujemną przerwą energetyczną

Pasmo przewodnictwa styka się tu w punkcie Γ z pasmem ciężkich dziur (mówimy, że pasma te są *zdegenerowane* w punkcie Γ). Oznacza to, że szerokość przerwy energetycznej jest równa zeru.

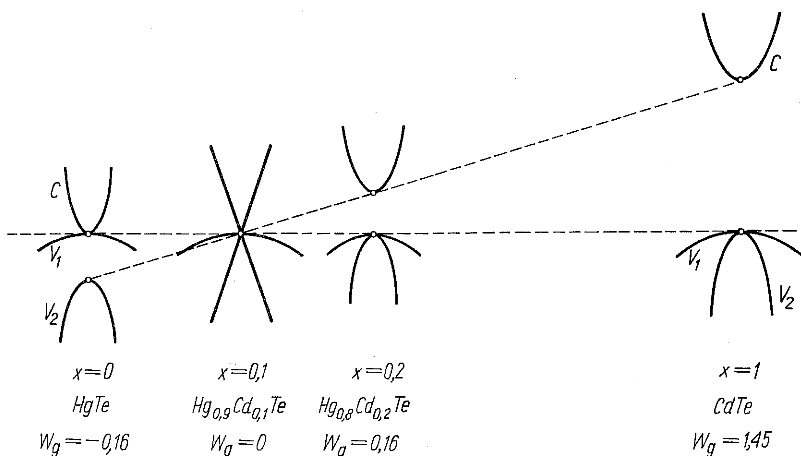
Ujemną przerwą energetyczną nazywa się natomiast różnicę energii między wierzchołkami pasm lekkich i ciężkich dziur. Wielkość ta, oznaczona na rysunku symbolem W_g, wynosi np. dla szarej cyny $-0,4$ eV, a dla HgTe $-0,16$ eV.

Mimo braku zwykłej przerwy energetycznej materiały te zachowują się jak typowe półprzewodniki. Nie wykazują one przewodnictwa elektrycznego w temperaturze zera bezwzględnego, ponieważ, podobnie jak w innych półprzewodnikach, pasmo walencyjne jest w tej temperaturze całkowicie zapełnione, a pasmo przewodnictwa – puste. Ze wzrostem temperatury ich przewodność wzrasta, co świadczy o tym, że koncentracja swobodnych nośników ładunku zwiększa się w miarę wzrostu energii drgań sieci krystalicznej[1].

W celu wyjaśnienia nazwy „półprzewodniki z ujemną przerwą energetyczną", rozważmy przebieg zmian struktury pasmowej kryształu mieszanego $Hg_{1-x}Cd_xTe$ w zależności od wartości współczynnika x – rys. 7/17.

Przy powiększaniu współczynnika x od zera obserwuje się początkowo zmniejszanie odległości między pasmami ciężkich i lekkich dziur, a przy $x = 0,1$ odległość ta staje się równa zeru. Po przekroczeniu wartości $x = 0,1$, otrzymujemy półprzewodnik z „normalną" (tzn. dodatnią) przerwą energetyczną. Jej szerokość wzrasta liniowo ze wzrostem x. W całym zakresie zmienności współczynnika x można szerokość przerwy energetycznej rozpatrywanego kryształu wyrazić wzorem

$$W_g = -0,16+1,61x \tag{7/16}$$



7/17 Struktury pasmowe kryształów mieszanych $Hg_{1-x}Cd_xTe$ o różnym składzie ($x = 0 \dots 1$)

[1] Koncentrację tę można wyznaczyć ze wzoru (8/27), podstawiając $W_g = 0$. W przeciwieństwie do innych półprzewodników, w których koncentracja nośników jest wykładniczą funkcją temperatury, w półprzewodnikach z zerową przerwą energetyczną jest ona proporcjonalna do $T^{3/2}$.

z tym, że dla $x < 0,1$, jako W_g przyjęto tu odległość między pasmami ciężkich i lekkich dziur.

Kryształy $Hg_{1-x}Cd_xTe$ o składzie odpowiadającym $x = 0,1$, wykazują specyficzne właściwości. Jak wynika z rys. 7/17 energia elektronu jest tu liniowo zależna od wartości wektora $\boldsymbol{k}$, a w punkcie $k = 0$ krzywizna funkcji $W(k)$ staje się nieskończenie wielka. Oznacza to, że masa efektywna elektronów dąży w tym punkcie do zera. W niskich temperaturach obserwuje się tu rzeczywiście niezwykle małe wartości masy efektywnej (poniżej $0,001m_0$), a w konsekwencji tego olbrzymie wartości ruchliwości nośników[1].

Ze względu na możliwość otrzymania dowolnej wartości przerwy energetycznej (w granicach od 0 do 1,45 eV) i ze względu na dużą wartość ruchliwości nośników, kryształy $Hg_{1-x}Cd_xTe$ znajdują zastosowanie jako detektory promieniowania podczerwonego.

7.6 DOMIESZKI W PÓŁPRZEWODNIKACH

7.6.1 Uwagi ogólne

Zasadniczą przyczyną, która sprawiła, że półprzewodniki zrobiły w ciągu ostatniego ćwierćwiecza tak oszałamiającą karierę, jest fakt niezwykłej zmienności ich własności fizycznych w wyniku wprowadzenia do nich drobnych ilości obcych atomów. Przez odpowiednie domieszkowanie można na przykład zmieniać przewodność elektryczną półprzewodnika w zakresie ponad 10 rzędów wielkości (zob. rys. 10/11), można uczynić go przezroczystym lub nieprzezroczystym dla promieniowania o danej długości fali, można wywołać jego świecenie, wytworzyć w odpowiednim miejscu kryształu barierę potencjału o pożądanej wysokości, zmienić własności magnetyczne kryształtu itd. Działanie olbrzymiej większości elementów półprzewodnikowych opiera się na korzystaniu z tych właśnie właściwości półprzewodników.

Wprowadzenie do kryształu obcych atomów zakłóca periodyczność sieci krystalicznej, a przez to powoduje zmianę struktury pasmowej półprzewodnika. W półprzewodniku pojawiają się dodatkowe stany kwantowe, którym odpowiadają energie leżące bądź w poszczególnych pasmach, bądź też w przerwie energetycznej. Te ostatnie mają, jak zobaczymy, największe znaczenie – one bowiem głównie są odpowiedzialne za modyfikację własności fizycznych półprzewodnika. Dodatkowe poziomy energetyczne, wytworzone wskutek wprowadzenia obcych atomów, nazywamy *poziomami domieszkowymi.*

W zależności od położenia danego poziomu domieszkowego w stosunku do krawędzi pasma walencyjnego i pasma przewodnictwa, dzieli się domieszki na 2 grupy o zasadniczo różnych własnościach. Rozróżnia się mianowicie domie-

---

[1] Z pojęciem ruchliwości nośników zapoznamy się w p. 10.3. W rozważanych kryształach uzyskiwano ruchliwości dochodzące do $3 \cdot 10^6$ cm²/V·s. Jest to wynik zaskakujący, ponieważ w najdoskonalszych monokryształach miedzi można w niskich temperaturach uzyskać ruchliwość rzędu 10^6 cm²/V·s, tu zaś w nieuporządkowanym krysztale mieszanym jest ona 3-krotnie większa.

szki *płytkie* i *głębokie*. Dla pierwszych poziomy domieszkowe znajdują się w niewielkiej odległości (kilka do kilkudziesięciu meV) poniżej dna pasma przewodnictwa lub powyżej wierzchołka pasma walencyjnego. Przy większych odległościach poziomów domieszkowych od krawędzi pasm mówi się natomiast o domieszkach głębokich.

Podział ten jest mało precyzyjny, nie bierze on bowiem pod uwagę różnic w fizycznej naturze poszczególnych rodzajów centrów domieszkowych. Dlatego też, w punkcie tym przy omawianiu własności różnych rodzajów domieszek posłużymy się nieco innym rozróżnieniem, podzielimy mianowicie domieszki na *wodoropodobne* i *zlokalizowane*.

7.6.2 Domieszki wodoropodobne

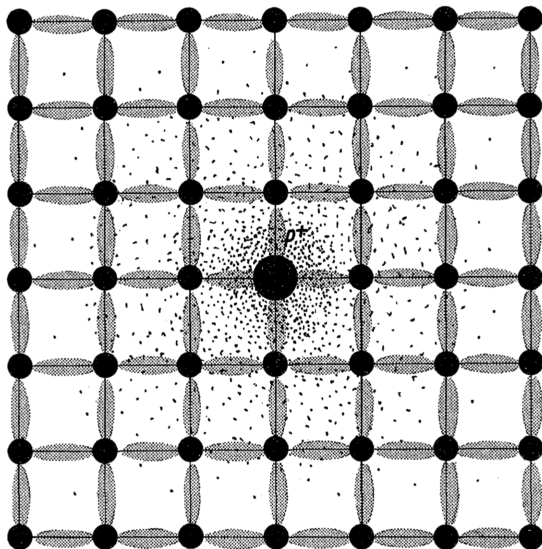
Są to domieszki, które z dość dużą dokładnością mogą być opisane za pomocą modelu atomu wodoropodobnego. Takimi domieszkami są atomy pierwiastków mających o 1 elektron walencyjny więcej lub o 1 elektron walencyjny mniej niż atomy własne półprzewodnika. W celu ustalenia uwagi, zajmijmy się jednym z najważniejszych, ze względu na zastosowania, półprzewodników, tj. krzemem. Atomy tego pierwiastka mają, jak pamiętamy, 4 elektrony walencyjne. Jednocześnie każdy atom ma w krysztale 4 najbliższych sąsiadów, z którymi jest połączony za pomocą wiązań kowalencyjnych – zob. rys. 5/9.

Jeżeli zamiast 1 atomu krzemu wprowadzimy w położenie węzłowe kryształu atom pierwiastka należącego do piątej grupy układu okresowego (np. atom fosforu), to 4 spośród jego 5 elektronów walencyjnych będą wykorzystane w wiązaniach krystalicznych, piąty natomiast „nadliczbowy" elektron będzie utrzymywany przy swoim macierzystym atomie siłami kulombowskimi. W rozważanym punkcie sieci krystalicznej będziemy więc mieli dodatnio naładowany jon P^+ i „krążący" dookoła niego elektron – zob. rys. 7/18. Jest to sytuacja analogiczna do tej, z którą mieliśmy do czynienia w atomie wodoru, dlatego też domieszkę tego rodzaju nazwaliśmy domieszką *wodoropodobną*.

Energia wiązania nadliczbowego elektronu w atomie fosforu wbudowanym w kryształ krzemu jest jednakże o kilka rzędów wielkości mniejsza niż energia wiązania elektronu w atomie wodoru. Można się o tym przekonać, analizując wzór (3/17), który tu dla wygody powtórzymy

$$W = -\frac{m_0 q^4}{32\pi^2 \varepsilon_0^2 \hbar^2} \tag{7/17}$$

Ośrodek krystaliczny wprowadza do tego wzoru 2 istotne poprawki. Po pierwsze, zamiast przenikalności elektrycznej próżni ε_0 należy tu podstawić $\varepsilon = \varepsilon_r \varepsilon_0$, gdzie ε_r jest stałą dielektryczną półprzewodnika (przykładowo dla krzemu $\varepsilon_r = 11{,}9$), po drugie zaś, zamiast masy elektronu swobodnego m_0 należy wpisać masę efektywną elektronu w krysztale. Ponieważ energia jonizacji atomu

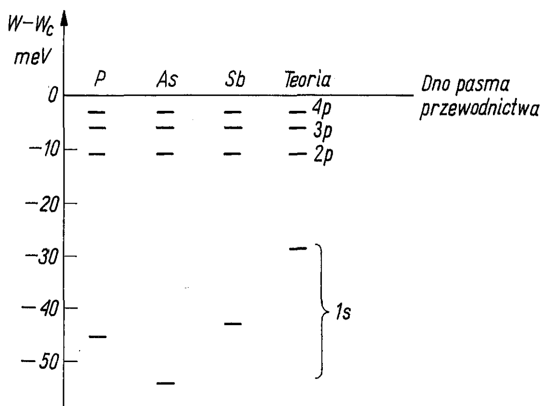


7/18 Poglądowy sposób przedstawienia donoru (P^+) wraz z należącym do niego elektronem — gęstość punktów reprezentuje gęstość prawdopodobieństwa znalezienia elektronu

wodoru wynosi 13,6 eV, zatem energia wiązania nadliczbowego elektronu z centrum domieszkowym może być wyznaczona ze wzoru

$$\Delta W_d = 13{,}6 \frac{m_n^*/m_0}{\varepsilon_r^2} \tag{7/18}$$

Podstawiając odpowiednie wartości parametrów materiałowych dla krzemu, otrzymuje się $\Delta W_d = 29$ meV. Rzeczywiste wartości energii wiązania odbiegają nieco od wyznaczonych ze wzoru (7/18). Jak wynika z danych zestawionych na rys. 7/19, są one nieco różne dla różnych pierwiastków. Wartości te zestawiono ponadto w tabl. 7/1.

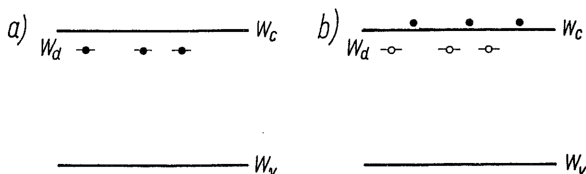


7/19 Poziomy energetyczne (poziom podstawowy i poziomy wzbudzone) dla różnych rodzajów donorów w krzemie

Na rysunku 7/19 pokazano nie tylko energię stanu podstawowego, lecz również energie kilku najniższych stanów wzbudzonych. Poziomem zerowym jest na tym rysunku dno pasma przewodnictwa, podobnie jak dla elektronu w atomie wodoru był nim poziom próżni. Elektron związany z centrum domieszkowym ma więc energię ujemną, elektron natomiast o energii dodatniej jest uwolniony z sił wiążących go z tym centrum i może się swobodnie poruszać w całym krysztale. Przynależy on wówczas do pasma przewodnictwa, a w centrum domieszkowym pozostaje nieruchomy, dodatnio zjonizowany atom domieszki.

Proces wzbudzenia elektronu z poziomu domieszkowego do pasma przewodnictwa nazywamy *jonizacją* atomu domieszkowego, a domieszkę, która w wyniku jonizacji oddaje elektron do pasma przewodnictwa nazywamy *donorem*[1].

Jonizacja domieszki donorowej może zachodzić na przykład pod wpływem energii drgań cieplnych kryształu. Proces ten pokazano na rys. 7/20. W temperaturze zera bezwzględnego elektrony obsadzają najniższe możliwe poziomy energetyczne. W tych warunkach wszystkie stany kwantowe odpowiadające poziomowi domieszkowemu (który w tym przypadku nazywamy *poziomem donorowym* W_d) są obsadzone, a więc atomy domieszek nie są zjonizowane.



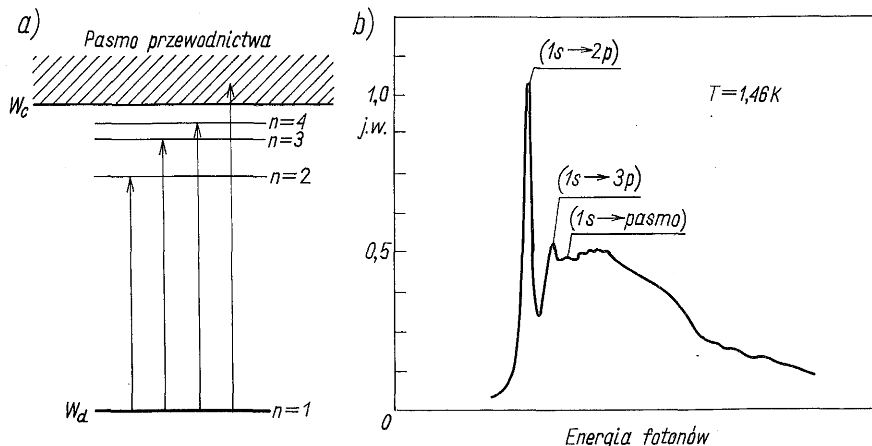
7/20 Obsadzenie poziomów donorowych w półprzewodniku w temperaturze $T = 0$ (a) i w temperaturze pokojowej (b)

W temperaturze pokojowej średnia energia drgań cieplnych przypadająca na 1 atom kryształu jest porównywalna z energią wiązania nadliczbowego elektronu z donorem. W wyniku tego, praktycznie rzecz biorąc, wszystkie donory są zjonizowane, a elektrony – wzbudzone do pasma przewodnictwa.

Wzbudzenie elektronów z poziomów donorowych może również zajść w wyniku działania promieniowania elektromagnetycznego. Ze względu na istnienie stanów wzbudzonych, obserwuje się przy tym charakterystyczne widmo absorpcyjne, będące analogiem serii Balmera dla wodoru. Przykład takiego widma pokazano na rys. 7/21b. W miarę wzrostu energii fotonów, elektrony są wzbudzane na coraz wyższe poziomy energetyczne (rys. 7/21a), a po osiągnięciu pewnej wartości energii (przy której $h\nu > \Delta W_d$) zachodzą fotojonizacyjne przejścia elektronów do pasma przewodnictwa.

Poza energią wiązania elektronu z domieszką donorową możemy na podstawie teorii atomu wodoru wyznaczyć najbardziej prawdopodobną odległość (w pm) tego elektronu od centrum domieszkowego. Oprzemy się w tym przypadku

[1] Nazwa ta pochodzi od łacińskiego słowa „dono" (darowuję, obdarzam).



7/21 Wzbudzenia elektronów z podstawowego poziomu donorowego (a) i towarzyszące im widmo absorpcyjne półprzewodnika (b)
(Opracowano na podstawie: G. E. Stillman, C. M. Wolfe, J. O. Dimmock: Magnetospectroscopy of shallow donors in GaAs. *Solid State Com.*, **7**, 921 (1969))

na wzorze (3/6), do którego zastosujemy omawiane poprzednio poprawki. W wyniku tego otrzymamy

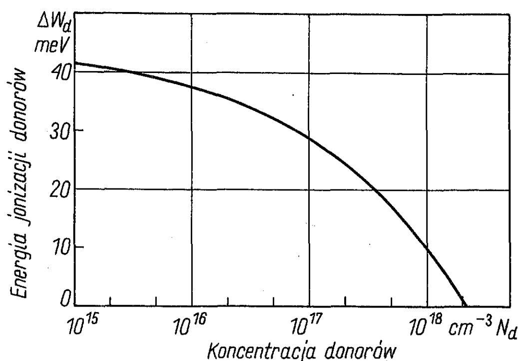
$$r_d = 53 \frac{\varepsilon_r}{m_n^*/m_0} \tag{7/19}$$

co po podstawieniu odpowiednich danych dla krzemu prowadzi do wniosku, że najbardziej prawdopodobna odległość nadliczbowego elektronu związanego z centrum donorowym wynosi ok. 2 nm. Ponieważ odległość między sąsiednimi atomami w krysztale krzemu jest równa 235 pm, zatem w kuli o promieniu 2 nm znajduje się kilka tysięcy atomów krzemu. Oznacza to, że nadliczbowy elektron związany z centrum donorowym jest w krysztale słabo zlokalizowany. Dla stanów wzbudzonych lokalizacja ta jest jeszcze słabsza, o czym można przekonać się, analizując wzór (3/5).

Przy małych koncentracjach donorów można przyjmować, że funkcje falowe nadliczbowych elektronów nie zachodzą na siebie. Wynika stąd, że elektrony te nie oddziałują na siebie, a odpowiadające im poziomy energetyczne nie ulegają rozszczepieniu. Są to więc poziomy dyskretne, tak jak to pokazano na rys. 7/20. Po przekroczeniu pewnej wartości koncentracji funkcje falowe zaczynają zachodzić na siebie. Proces ten rozpoczyna się od wyższych stanów wzbudzonych, dla których funkcje falowe, jak to już zaznaczono, rozciągają się na większą odległość od centrum donorowego. Odpowiednie poziomy energetyczne rozszczepiają się więc w pasma, których szerokość wzrasta w miarę powiększania koncentracji donorów. Pasma te zlewają się stopniowo z pasmem przewodnictwa, co uwidacznia się zanikiem kolejnych linii absorpcyjnych w widmie przedstawionym na rys. 7/21.

Problem. 7.2. Oszacuj wartość koncentracji donorów w krzemie, po przekroczeniu której podstawowy poziom donorowy zaczyna rozszczepiać się w pasmo domieszkowe.

Zlewanie się kolejnych pasm domieszkowych z pasmem przewodnictwa można interpretować jako stopniowe obniżanie się dna tego pasma, z czym wiąże się zmniejszanie efektywnej energii jonizacji donorów. Proces stopniowego zmniejszania się ΔW_d w miarę powiększania koncentracji donorów N_d zilustrowano na rys. 7/22. Po przekroczeniu pewnej wartości N_d, która, jak widzimy, dla krzemu wynosi $2 \cdot 10^{18}$ cm^{-3}, pasma domieszkowe są całkowicie wchłonięte przez pasmo przewodnictwa. W pasmie przewodnictwa występują więc swobodne nośniki ładunku (elektrony), których koncentracja, podobnie jak to występuje w metalu, nie zależy od temperatury. Bardzo silnie domieszkowane półprzewodniki mają więc własności zbliżone do metali.

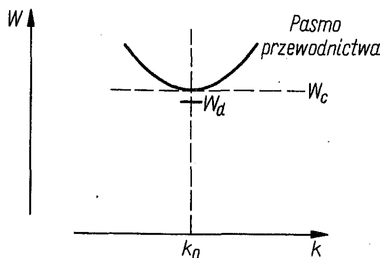


7/22 Zależność napięcia jonizacji donoru ΔW_d od koncentracji donorów N_d dla krzemu domieszkowanego fosforem

Obok lokalizacji nadliczbowego elektronu (związanego z centrum donorowym) w przestrzeni rzeczywistej może nas interesować lokalizacja podstawowego poziomu donorowego w przestrzeni k. Można ją wyznaczyć, korzystając z odpowiedniej relacji Heisenberga. Opierając się na wzorze (1/38) i oznaczając przez Δx nieoznaczoność położenia elektronu, a przez Δk – nieoznaczoność jego wektora falowego, dochodzimy do wniosku, że iloczyn tych 2 wielkości jest rzędu jedności

$$\Delta x \cdot \Delta k \approx 1 \qquad (7/20)$$

Przyjmując dla krzemu, zgodnie z poprzednimi obliczeniami, $\Delta x \approx 2$ nm, otrzymujemy $\Delta k \approx 5 \cdot 10^6$ cm^{-1}, co stanowi ok. 5% odpowiedniego promienia strefy Brillouina. Centrum donorowe jest więc słabo zlokalizowane w przestrzeni rzeczywistej, natomiast dość dobrze zlokalizowane w przestrzeni wektora falowego – zob. rys. 7/23.



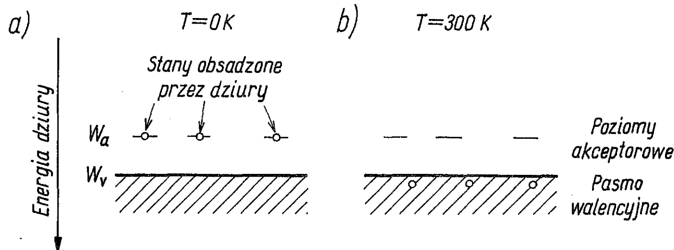
7/23
Poziom donorowy na wykresie W(k)

Zajmijmy się teraz drugim rodzajem domieszek wodoropodobnych, tzn. atomami pierwiastków mających o 1 elektron walencyjny mniej niż atomy własne półprzewodnika. Dla krzemu będą to pierwiastki trójwartościowe, spośród których największe znaczenie ma bor.

Atom boru wbudowany w miejscu węzłowym kryształu krzemu ma o 1 elektron za mało, aby w pełni wysycić wiązania tetraedryczne występujące w tej strukturze krystalicznej. Korzystając z wprowadzonego w poprzednim rozdziale pojęcia dziury, można w bardzo prosty sposób opisać własności tych domieszek. W atomie boru nic się bowiem nie zmieni, jeżeli dodamy mu czwarty elektron walencyjny oraz dziurę, która zgodnie z definicją ma taką masę i taki ładunek, że suma mas i ładunków dwóch dodatkowo wprowadzonych cząstek jest równa zeru.

Dodanie do atomu boru czwartego elektronu walencyjnego przekształca go w jon B^-. W polu elektrycznym tego jonu porusza się dziura o dodatnim ładunku elektrycznym, a zatem sytuacja jest, z wyjątkiem znaków ładunków, analogiczna do rozpatrywanej poprzednio. Omawiany tu sposób postępowania nazywany modelem *odwróconego* atomu wodoropodobnego. Do takiego atomu stosują się wszystkie poprzednie rozważania oraz wyprowadzone wzory z tym tylko zastrzeżeniem, że zamiast masy efektywnej elektronu w pasmie przewodnictwa należy w nich wpisać masę efektywną dziury w pasmie walencyjnym.

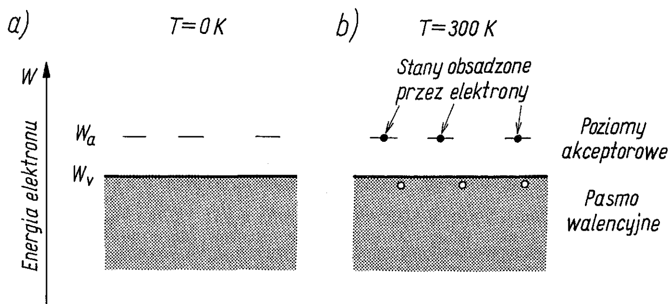
Drugą sprawą, o której należy pamiętać, jest to, że w odwróconym atomie wodoropodobnym obowiązuje odwrócony kierunek osi energii — zob. rys. 7/24.



7/24 Obsadzenie poziomów akceptorowych przez dziury w temperaturze $T = 0$ (a) i w temperaturze pokojowej (b)

W temperaturze zera bezwzględnego dziury obsadzają stany kwantowe o najniższych energiach, tzn. stany domieszkowe. Atomy domieszkowe są wówczas elektrycznie neutralne, czyli nie są zjonizowane. Pod wpływem energii cieplnej (lub energii promieniowania elektromagnetycznego) zachodzi wzbudzenie dziur z poziomów domieszkowych do pasma walencyjnego, przy czym w temperaturze pokojowej, praktycznie rzecz biorąc, wszystkie dziury są wzbudzone do pasma walencyjnego (rys 7/24b). W centrach domieszkowych znajdują się wówczas nieruchome, ujemnie zjonizowane atomy domieszek.

Zachowanie się trójwartościowych domieszek w krysztale krzemu można również opisać nie posługując się pojęciem dziury. Ten sposób postępowania pokazano na rys. 7/25. W temperaturze zera bezwzględnego elektrony walencyjne półprzewodnika obsadzają wszystkie stany kwantowe w pasmie walencyjnym,



7/25 Obsadzenie poziomów akceptorowych przez elektrony w temperaturze $T = 0$ (a) w temperaturze pokojowej (b)

pozostawiając nie obsadzone poziomy domieszkowe. W centrach domieszkowych znajdują się wówczas elektrycznie obojętne atomy boru. Pod wpływem energii drgań cieplnych kryształu następuje wzbudzenie elektronów z pasma walencyjnego na poziomy domieszkowe. W pasmie walencyjnym pojawiają się zatem nieobsadzone stany kwantowe, a w centrach domieszkowych – zjonizowane atomy domieszek – zob. rys. 7/25b.

Proces jonizacji trójwartościowego atomu domieszkowego polega zatem na przyjęciu elektronu z pasma walencyjnego. Z tego względu domieszkę, która w wyniku jonizacji staje się ujemnie naładowanym jonem, nazywamy *akceptorem*[1]. Wartości energii jonizacji najczęściej spotykanych donorów i akceptorów podano w tabl. 7/1.

7/1 Energie jonizacji donorów i akceptorów w krzemie i w germanie [meV]

Materiał podstawowy		Si	Ge	Materiał podstawowy		Si	Ge
Donory:	As	54	14	Akceptory:	Al	67	11
	Bi	69	13		B	45	11
	P	45	13		Ga	74	11
	Sb	43	10		In	153	12

Źródło: Landolt-Börnstein: Numerical data and functional relationships in science and technology. Vol. **17**: Semiconductors, Subvolume **a**. Springer-Verlag, Berlin 1982.

Rozważania te odnoszą się zarówno do krzemu, jak i do germanu. W tym ostatnim, ze względu na inne wartości parametrów materiałowych, uzyskuje się znacznie mniejsze wartości energii jonizacji domieszek – zob. tabl. 7/1.

Nieco inaczej wygląda sprawa w przypadku takich półprzewodników, jak np. arsenek galu. Ponieważ gal jest pierwiastkiem 3-wartościowym, a arsen 5-wartościowym, zatem donorami w GaAs mogą być atomy pierwiastków 6-wartościowych (np. siarki lub selenu) wbudowane w miejsce arsenu, akceptorami zaś – atomy pierwiastków 2-wartościowych (np. cynku) wbudowane w miejsce galu. Atomy pierwiastków należących do IV grupy układu

[1] Nazwa pochodzi od łacińskiego słowa „accepto" (przyjmuję).

okresowego (np. krzemu lub germanu) mogą być zarówno donorami, jak i akceptorami. Atom taki znajdujący się w miejscu atomu galu pełni rolę donora, natomiast w miejscu atomu arsenu — akceptora. Energie jonizacji donorów w GaAs zawierają się w granicach 5,75–5,93 meV, akceptorów zaś – w granicach 27–40 meV.

7.6.3* Domieszki zlokalizowane

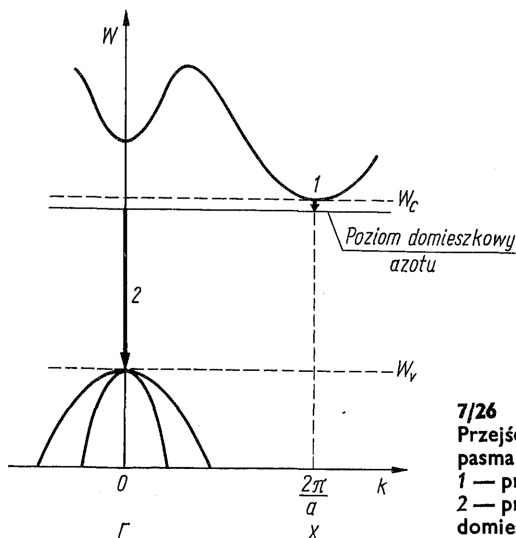
W poprzednim punkcie widzieliśmy, że domieszki wodoropodobne są w krysztale słabo zlokalizowane. W wyniku tego odpowiadające im poziomy energetyczne są dość dobrze zlokalizowane w przestrzeni *k*. Oznacza to, że nieoznaczoność wektora falowego jest mała w stosunku do promienia strefy Brillouina. W przeciwieństwie do tego, istnieją domieszki, które z różnych przyczyn zaburzają jedynie niewielki obszar kryształu (o promieniu rzędu 0,1–0,2 nm), są one więc dobrze zlokalizowane w przestrzeni. Konsekwencją tego faktu jest duża nieoznaczoność wektora falowego. Korzystając z zależności (7/20), można stwierdzić, że obejmuje ona całą strefę Brillouina. Poziom energetyczny takiej domieszki należy więc na wykresie przedstawiającym strukturę pasmową półprzewodnika rysować nie w postaci krótkiego odcinka, jak to występowało dla domieszki wodoropodobnej (zob. rys. 7/23), lecz w postaci linii poziomej rozciągającej się na całą strefę Brillouina – rys. 7/26.

Jako przykład domieszki zlokalizowanej rozpatrzmy azot w fosforku galu (GaP). Półprzewodnik ten ma skośną przerwę energetyczną z minimum pasma przewodnictwa w punkcie *X*. Ponieważ przejścia międzypasmowe w czystym GaP są możliwe tylko przy udziale fononów, intensywność ich jest tak mała, że dioda elektroluminescencyjna wykonana z czystego GaP, w praktyce nie świeci.

Jeżeli teraz w miejsce atomu fosforu wprowadzimy atom należący do tej samej grupy układu okresowego (np. azot), to pod względem wiązań krystalicznych będzie się on zachowywał identycznie jak atom fosforu. Pierwiastek tego rodzaju nazywamy domieszką *izoelektronową*, co oznacza, że atom domieszkowy ma taką samą liczbę elektronów walencyjnych co atom własny, którego miejsce zajmuje w krysztale.

Zastąpienie atomu fosforu atomem azotu nie pozostaje jednak bez konsekwencji. Rdzenie tych atomów różnią się bowiem zarówno rozmiarami, jak i rozkładem potencjału. W wyniku tego wytwarza się w krysztale, silnie zlokalizowana studnia potencjału, która może wiązać elektron. Atomy azotu wytwarzają więc w GaP poziom domieszkowy, leżący w odległości ok. 10 meV pod dnem pasma przewodnictwa – zob. rys. 7/26. Poziom ten odgrywa istotną rolę w procesie rekombinacji nośników (i związanego z nią świecenia rekombinacyjnego) w rozpatrywanym półprzewodniku.

Przejście elektronu z pasma przewodnictwa do pasma walencyjnego zachodzi tu w sposób dwustopniowy. W pierwszej fazie elektron spada z pasma przewodnictwa na poziom domieszkowy (przejście bezpromieniste), oddając przy tym posiadany pęd domieszce. Stąd następuje przejście promieniste do pasma walencyjnego. Ze względu na dużą nieoznaczoność wektora falowego, można w uproszczonym rozumowaniu uważać, że jest to przejście proste zachodzące w punkcie

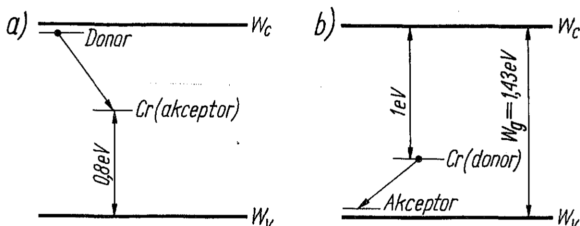


7/26
Przejście elektronu z pasma przewodnictwa do pasma walencyjnego w GaP
1 — przejście na poziom domieszkowy;
2 — przejście promieniste z poziomu domieszkowego

Γ. Dzięki temu, że prawdopodobieństwo takich przejść jest duże, wydajność świetlna diod wykonanych z fosforku galu domieszkowanego azotem jest o kilka rzędów wielkości większa niż diod z czystego GaP.

Innym przykładem domieszki zlokalizowanej, odgrywającej ważną rolę w elementach półprzewodnikowych, jest chrom w arsenku galu. Atom chromu ma konfigurację elektronową $3d^5 4s^1$. Wbudowując się w krysztale GaAs w miejsce galu, oddaje on 3 elektrony do wiązań krystalicznych i przechodzi do stanu o konfiguracji $3d^3$. Właściwości takiego centrum domieszkowego są dość nieoczekiwane. Jako domieszka wieloelektronowa chrom może znajdować się w różnych stanach ładunkowych i dzięki temu może być zarówno głębokim akceptorem, jak i głębokim donorem. Rozpatrzmy kolejno te 2 poziomy domieszkowe.

Jeżeli w arsenku galu oprócz chromu występują płytkie donory (rys. 7/27a), to nadliczbowe elektrony tych donorów nie będą przechodziły do pasma przewodnictwa, lecz będą „spadały" na położony niżej poziom akceptorowy. Pasmo przewodnictwa pozostanie więc prawie puste. Mówimy wówczas, że działanie donorów zostało skompensowane działaniem domieszki chromowej, lub też, że domieszka ta stanowi *pułapkę* dla elektronów wprowadzonych przez donory.



7/27 Działanie domieszki chromowej w GaAs zawierającym płytkie donory (a) lub płytkie akceptory (b)

Jeśli natomiast w arsenku galu występują płytkie akceptory (rys. 7/27b), to działanie ich zostanie również skompensowane wskutek istnienia wprowadzonego przez chrom poziomu donorowego. W tym przypadku dziury z poziomów akceptorowych nie będą przechodziły do pasma walencyjnego, lecz będą „spadały" na położony niżej (dla dziur!) poziom domieszkowy. Inaczej mówiąc, atomy chromu oddadzą po jednym ze swych elektronów akceptorom (rys. 7/27b), powodując ich jonizację, a przez to uniemożliwiając im przyjmowanie elektronów z pasma walencyjnego. Pasmo to pozostanie więc również prawie puste.

Tak więc dzięki zastosowaniu domieszki chromowej można otrzymać arsenek galu prawie całkowicie pozbawiony nośników ładunku. Jest to tzw. *półizolacyjny* GaAs, charakteryzujący się oporem właściwym rzędu $10^8 \Omega \cdot cm$. Jest on znakomitym materiałem podłożowym wykorzystywanym w pewnych typach elementów półprzewodnikowych i układów scalonych.

Bez domieszki chromowej otrzymalibyśmy materiał niskooporowy, ponieważ w procesie hodowania kryształów GaAs nie daje się uniknąć wprowadzenia niepożądanych domieszek, przede wszystkim krzemu, który zajmując miejsca galu wytwarza płytkie poziomy donorowe.

Innym sposobem otrzymywania półizolacyjnego GaAs jest hodowanie kryształu w atmosferze, w której występuje nadmiar arsenu. Atomy arsenu wbudowują się wówczas częściowo w miejsca atomów galu, wskutek czego powstają wspomniane już defekty określane w literaturze jako „antysite defects". Defekt ten jest głębokim donorem, któremu odpowiada poziom energetyczny znajdujący się w odległości 0,8 eV poniżej dna pasma przewodnictwa. Spełnia on w stosunku do płytkich akceptorów rolę kompensującą w sposób podobny do przedstawionego na rys. 7/27b.

Z innymi jeszcze rodzajami domieszek, odgrywającymi ważną rolę w półprzewodnikach, zapoznamy się w p. 9.3.1.

8

KONCENTRACJE NOŚNIKÓW ŁADUNKU W PÓŁPRZEWODNIKU W STANIE RÓWNOWAGI TERMODYNAMICZNEJ

Przy rozpatrywaniu działania elementów półprzewodnikowych zasadnicze znaczenie ma znajomość rozkładu ładunków elektrycznych zgromadzonych w poszczególnych obszarach danego elementu oraz prądów przepływających przez te obszary. Wielkości te zależą od wypadkowego działania olbrzymiej ilości nośników ładunku (elektronów i dziur). Przy analizowaniu zjawisk elektronowych w półprzewodnikach nie interesujemy się jednakże zachowaniem każdego pojedynczego nośnika, lecz pewnymi uśrednionymi wielkościami charakteryzującymi cały zbiór tych nośników w określonym obszarze.

Jedną z takich wielkości jest *koncentracja* nośników danego rodzaju, tzn. liczba nośników przypadająca na jednostkę objętości półprzewodnika. W rozdziale tym zajmiemy się koncentracjami elektronów i dziur w półprzewodniku w warunkach, które określamy jako *stan równowagi termodynamicznej*.

8.1 RÓWNOWAGA TERMODYNAMICZNA

Każdy układ fizyczny znajdujący się w niezmiennych w czasie warunkach zewnętrznych i nie oddziałujący z otoczeniem osiąga po dostatecznie długim czasie stan równowagi, który nazywamy *równowagą termodynamiczną*. Charakteryzuje się ona następującymi dwiema cechami:

1) stan układu jest *ustalony*, tzn. opisujące go parametry makroskopowe (takie np. jak temperatura, energia, koncentracja nośników itd.) nie zmieniają się w czasie.

2) nie występują żadne *przepływy* (np. energii, nośników itp.) ani wewnątrz układu, ani też między tym układem a otoczeniem.

Mówiąc o tym, że parametry stanu układu są niezależne od czasu, mamy tu na myśli wartości uśrednione w kolejnych dostatecznie długich odcinkach czasu.

Wartości chwilowe wielkości opisujących stan układu nie są na ogół równe tym wartościom średnim, lecz wykazują nieuporządkowane, przypadkowe odchylenia od wartości średnich. Odchylenia te, zwane *fluktuacjami* są wynikiem atomowej struktury materii[1]. Przy dużej liczbie nośników są one niewielkie – relatywnie tym mniejsze, im większa jest koncentracja nośników. Przykładowo, przy koncentracji równej 10^{10} elektronów na cm^3 średnie odchylenie kwadratowe wynosi $10^5 \, cm^{-3}$, co stanowi zaledwie jedną tysięczną procentu.

Podobnie należy rozumieć stwierdzenie, że w stanie równowagi termodynamicznej nie występują żadne przepływy. Nie oznacza to bynajmniej, że równowaga ta jest stanem całkowitego bezruchu. Pod względem mikroskopowym mamy tu bowiem do czynienia z nieustannie zachodzącymi procesami przepływu energii, cząstek itd. Jednakże, jeżeli w stanie równowagi termodynamicznej zachodzi przepływ cząstek danego rodzaju (np. elektronów lub fotonów) z obszaru A do obszaru B, to jednocześnie zachodzi proces odwrotny, a więc przepływ takich samych cząstek w przeciwnym kierunku, przy czym oba te przepływy zachodzą z taką samą intensywnością, tzn., *statystycznie* rzecz biorąc, równoważą się wzajemnie[2].

W warunkach równowagi termodynamicznej średnie wartości natężenia strumieni cząstek przepływających w obu kierunkach są więc równe, co oznacza, że w sensie makroskopowym przepływ cząstek nie występuje.

Podobnie, jeżeli np. elektrony są wzbudzane z poziomu energetycznego W_1, na wyższy poziom W_2, to jednocześnie, statystycznie rzecz biorąc, taka sama liczba elektronów przechodzi z poziomu W_2 na poziom W_1.

Stwierdzenia te są wyrazem tzw. *zasady równowagi szczegółowej*, która orzeka, że w warunkach równowagi termodynamicznej każdy możliwy do wyodrębnienia proces fizyczny jest kompensowany przez swój proces odwrotny. Kompensacja ta zachodzi przy tym niezależnie od innych równocześnie zachodzących procesów.

W półprzewodniku zachodzą nieustannie procesy generacji i rekombinacji par elektron-dziura, a niezależnie od nich przejścia elektronów z poziomów domieszkowych do pasm energetycznych i w odwrotnych kierunkach. Zgodnie z zasadą równowagi szczegółowej szybkość generacji par elektron-dziura jest dokładnie równa szybkości rekombinacji tych par, niezależnie od innych procesów, w wyniku których powstają i zanikają nośniki. Wskutek równoważenia się wszystkich wymienionych procesów ustalają się w danej temperaturze określone koncentracje elektronów i dziur.

---

[1] Fluktuacje te są np. przyczyną szumów elementów półprzewodnikowych.

[2] Słowa „statystycznie rzecz biorąc" oznaczają, że przy porównywaniu liczb cząstek przepływających w różnych kierunkach należy brać pod uwagę dostatecznie długie odcinki czasu.

8.2 STATYSTYKA FERMIEGO-DIRACA

Badaniem układów złożonych z dużej liczby cząstek zajmuje się *fizyka statystyczna*. W zależności od tego, czy do opisu zachowania się cząstek stosują się prawa mechaniki klasycznej, czy też mechaniki kwantowej, rozróżnia się statystykę klasyczną i statystyki kwantowe.

W mechanice klasycznej cząstki mają określoną indywidualność. Oznacza to, że można je np. w chwili początkowej „ponumerować", a następnie śledzić ruch każdej cząstki po jej torze i dzięki temu rozpoznać ją w dowolnej późniejszej chwili.

W opisie kwantowomechanicznym, w którym pojęcie toru cząstki traci sens, operuje się prawdopodobieństwem znalezienia cząstki w określonym obszarze. Fakt ten wyklucza możliwość śledzenia cząstki, a tym samym również możliwość rozróżniania jednakowych cząstek[1]. W tych warunkach nie można więc zastanawiać się nad tym, które cząstki znajdują się w danym stanie, a jedynie nad tym, ile cząstek znajduje się w tym stanie.

W przypadku elektronów, zgodnie z zakazem Pauliego, w każdym stanie kwantowym może znajdować się co najwyżej 1 elektron. Stan kwantowy może więc być albo obsadzony, albo też nie obsadzony przez elektron. W warunkach równowagi termodynamicznej ustala się w danym układzie fizycznym określony rozkład obsadzeń, opisany tzw. *funkcją rozkładu Fermiego-Diraca*, określającą prawdopodobieństwo obsadzenia przez elektrony stanów o energii W. Funkcja ta ma następującą postać:

$$f(W) = \frac{1}{1+\exp\dfrac{W-W_F}{kT}} \tag{8/1}$$

przy czym W_F oznacza tzw. *energię Fermiego*, nazywaną także *poziomem Fermiego*. Dla $W = W_F$ otrzymujemy $f(W) = 0{,}5$, a zatem energia Fermiego jest to wartość energii, dla której prawdopodobieństwo obsadzenia stanów wynosi 50%.

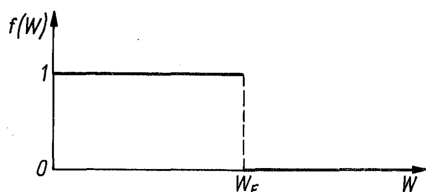
Jak wynika ze wzoru (8/1), kształt funkcji $f(W)$ zależy od temperatury, która w warunkach równowagi termodynamicznej jest oczywiście jednakowa w całym układzie.

Dla $T \to 0$ funkcja rozkładu staje się funkcją skokową pokazaną na rys. 8/1. Jak widzimy, wszystkie stany o energiach mniejszych niż W_F są obsadzone przez elektrony, natomiast wszystkie stany o energiach większych niż W_F pozostają nie obsadzone.

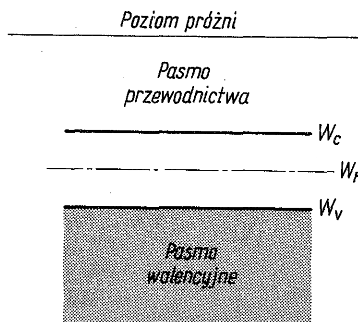
Obsadzenie wszystkich poziomów energetycznych aż do pewnego poziomu granicznego W_F jest uwarunkowane z jednej strony tendencją każdego układu fizycznego do przyjęcia takiego stanu, w którym energia układu osiąga wartość minimalną, z drugiej zaś zakazem Pauliego, zgodnie z którym stan kwantowy nie może być obsadzony przez więcej niż 1 elektron. Sytuacja jest tu w pewnym

---

[1] Stwierdzenie to jest treścią tzw. *zasady nierozróżnialności* – jednego z podstawowych praw fizyki kwantowej.



8/1 Funkcja rozkładu Fermiego-Diraca dla $T \to 0$



8/2 Położenie poziomu Fermiego w półprzewodniku samoistnym w temperaturze $T = 0$

sensie podobna do występującej w naczyniu wypełnionym cząsteczkami wody aż do pewnej wysokości, a więc również do pewnej wartości energii potencjalnej. W tym przypadku odpowiednikiem zakazu Pauliego jest zakaz zajmowania tego samego miejsca przez więcej niż jedną cząsteczkę wody.

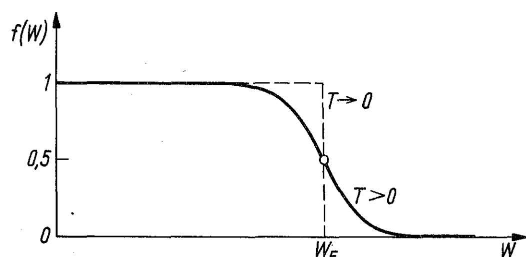
W metalach najwyższe pasmo energetyczne jest tylko częściowo obsadzone przez elektrony. Wypływa stąd oczywisty wniosek, że poziom Fermiego znajduje się wewnątrz tego pasma, tak jak to pokazano na rys. 6/3.

W półprzewodniku samoistnym w temperaturze $T = 0$ pasmo walencyjne jest całkowicie obsadzone przez elektrony, a pasmo przewodnictwa – całkowicie puste. Poziom Fermiego znajduje się więc w przerwie energetycznej i, jak dalej zobaczymy, leży prawie dokładnie w środku tej przerwy – rys. 8/2.

W temperaturach $T > 0$ funkcja rozkładu Fermiego-Diraca ma przebieg pokazany na rys. 8/3. Jak widzimy, część stanów kwantowych o energiach mniejszych niż W_F jest nie obsadzona przez elektrony, natomiast część elektronów, dzięki dostarczonej do układu energii cieplnej, obsadza stany o energiach większych niż W_F.

Krzywa przedstawiona na rys. 8/3 jest symetryczna względem punktu $(W_F, 1/2)$. Oznacza to, że prawdopodobieństwo obsadzenia przez elektron stanu o energii $W_F + \Delta W$ jest równe prawdopodobieństwu nieobsadzenia stanu o energii $W_F - \Delta W$ dla dowolnej wartości ΔW. Tę własność funkcji rozkładu można zapisać następująco:

$$1 - f\left(\frac{W - W_F}{kT}\right) = f\left(\frac{W_F - W}{kT}\right) \tag{8/2}$$



8/3 Funkcja rozkładu Fermiego-Diraca

Wzorem tym będziemy się posługiwali przy obliczaniu koncentracji dziur w półprzewodniku. Wynika z niego bowiem, że funkcja Fermiego-Diraca może być stosowana nie tylko do wyznaczania prawdopodobieństwa obsadzenia stanu w pasmie przewodnictwa przez elektron, lecz również do wyznaczania prawdopodobieństwa obsadzenia stanu w pasmie walencyjnym przez dziurę, z tym, że w tym ostatnim przypadku należy zmienić zwrot osi energii.

Problem 8.1. Wykaż słuszność wzoru (8/2).

W dostatecznie dużych odległościach od poziomu Fermiego można funkcję rozkładu wyrazić w przybliżeniu za pomocą prostszej zależności. Dla $W - W_F \gg kT$ można bowiem w wyrażeniu (8/1) pominąć w mianowniku jedność wobec składnika wykładniczego. Sprowadza to statystykę Fermiego-Diraca do statystyki klasycznej Maxwella-Boltzmanna (mówimy wówczas o tzw. *przybliżeniu boltzmannowskim*). Funkcja rozkładu wyraża się wówczas wzorem

$$f(W) = \exp\left(-\frac{W - W_F}{kT}\right) \tag{8/3}$$

Z tej przybliżonej postaci funkcji $f(W)$ będziemy często korzystali, ponieważ upraszcza ona znacznie obliczenia.

8.3 KONCENTRACJE RÓWNOWAGOWE — ZALEŻNOŚCI PODSTAWOWE

Przystąpimy teraz do zasadniczego tematu tego rozdziału, tj. do wyznaczenia koncentracji nośników (elektronów i dziur) w półprzewodniku w warunkach równowagi termodynamicznej. Koncentracje te, nazywane *koncentracjami równowagowymi* lub *normalnymi*, oznaczane są odpowiednio symbolami n_0 i p_0. Rozważania zawarte w tym punkcie odnoszą się zarówno do półprzewodników samoistnych, jak i domieszkowych. Zakładamy tu jedynie, że półprzewodnik jest *jednorodny*, tzn. jego skład chemiczny jest jednakowy w całej objętości. Założenie to dotyczy całego rozdziału.

W celu wyznaczenia koncentracji elektronów w pasmie przewodnictwa należy rozważyć, ile stanów kwantowych zawiera to pasmo i jakie jest prawdopodobieństwo obsadzenia poszczególnych stanów. Liczbę stanów przypadających na przedział energii od W do $W + \mathrm{d}W$ i na jednostkę objętości półprzewodnika wyrażamy jako $g(W)\,\mathrm{d}W$, przy czym wielkość $g(W)$ nazywamy *gęstością stanów*.

Prawdopodobieństwo obsadzenia stanu o energii W przez elektron wyraża poznana w poprzednim punkcie funkcja Fermiego-Diraca $f(W)$. Koncentracja elektronów o energiach zawartych w granicach od W do $W + \mathrm{d}W$ wynosi więc

$$\mathrm{d}n = f(W)\,g(W)\,\mathrm{d}W \tag{8/4}$$

W celu wyznaczenia całkowitej koncentracji elektronów należy wyrażenie to scałkować w przedziale energii stanowiącym pasmo przewodnictwa. A zatem

$$n_0 = \int_{W_c}^{\infty} f(W) g(W) \, dW \tag{8/5}$$

Wykładniczy charakter funkcji $f(W)$ sprawia, że iloczyn $f(W)g(W)$ staje się pomijalnie mały w niewielkiej odległości od dna pasma przewodnictwa. Z tego powodu górna granica całkowania we wzorze (8/5) nie ma istotnego znaczenia. W celu ułatwienia obliczeń przyjęto tu, że całkowanie jest rozciągnięte do nieskończoności.

Dla półprzewodnika z prostą przerwą energetyczną, w którym pasmo przewodnictwa wykazuje w pobliżu swego minimum symetrię sferyczną (a więc np. dla GaAs), gęstość stanów w pobliżu dna pasma przewodnictwa wynosi[1]

$$g(W) = 4\pi \left(\frac{2m_n^*}{h^2} \right)^{3/2} (W - W_c)^{1/2} \tag{8/6}$$

wzrasta ona więc proporcjonalnie do pierwiastka z energii mierzonej od dna pasma.

Dla półprzewodników nie spełniających podanych tu warunków (a więc np. dla krzemu i germanu) można gęstość stanów w pasmie przewodnictwa wyrazić formalnie za pomocą wzoru (8/6) z tym tylko, że zamiast masy efektywnej m_n^* należy podstawić pewną uśrednioną masę m_c zwaną *masą efektywną gęstości stanów* dla pasma przewodnictwa. Tak skorygowany wzór (8/6) przybiera postać

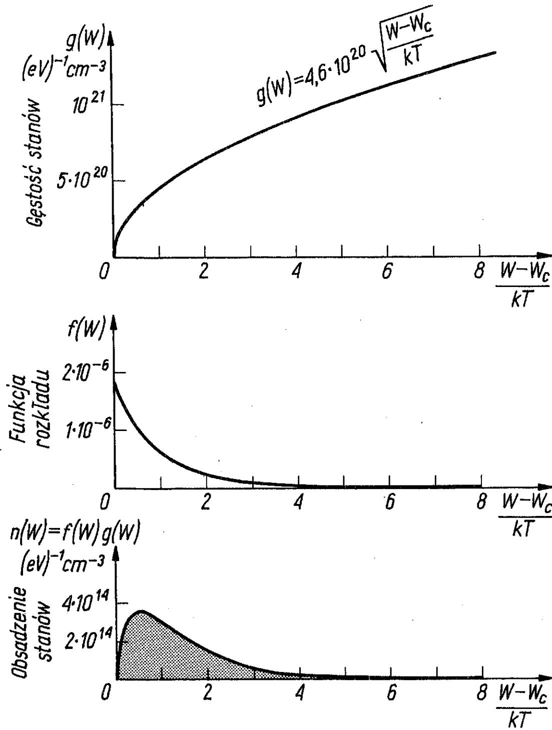
$$g(W) = 4\pi \left(\frac{2m_c}{h^2} \right)^{3/2} (W - W_c)^{1/2} \tag{8/7}$$

Przykładowy przebieg funkcji $f(W)$ i $g(W)$ pokazano na rys. 8/4. Na osi odciętych odłożono tu energię liczoną od dna pasma przewodnictwa wyrażoną w jednostkach kT.

Na rysunku przedstawiono również wykres funkcji $n(W) = f(W) g(W)$. Jak wynika z definicji funkcji $g(W)$, wyrażenie $n(W) \, dW$ oznacza koncentrację elektronów o energiach zawartych w przedziale od W do $W + dW$. Z tego względu wielkość $n(W)$ nazywamy *obsadzeniem stanów* w pasmie przewodnictwa.

Z rysunku 8/4 wynika, że tak jak to już wyżej zaznaczono, w paśmie tym są obsadzone stany o energiach (mierzonych od dna tego pasma) zawartych w granicach kilku jednostek kT. Wyższe poziomy energetyczne, praktycznie biorąc, są nie obsadzone. Zmiana położenia poziomu Fermiego zmienia obsadzenie stanów, nie naruszając jednakże zakresu energii, w którym obsadzenie to jest istotnie różne od zera.

[1] Wzór ten jest słuszny dla półprzewodników o niezbyt dużej koncentracji domieszek. Zjawiska występujące przy dużych koncentracjach domieszek będą omawiane w p. 8.7.



8/4
Przebieg funkcji $g(W)$, $f(W)$ oraz $n(W) = f(W)\,g(W)$ dla pasma przewodnictwa germanu samoistnego w temperaturze $T = 300$ K

Podstawiając odpowiednie wyrażenia na $f(W)$ i $g(W)$ do wzoru (8/5), otrzymujemy

$$n_0 = 4\pi\left(\frac{2m_c}{h^2}\right)^{3/2} \int\limits_{W_c}^{\infty} \frac{(W-W_c)^{1/2}\,\mathrm{d}W}{1+\exp\dfrac{W-W_F}{kT}} \tag{8/8}$$

Zapis ten można uprościć, wprowadzając następujące energie zredukowane:

$$x = \frac{W-W_c}{kT} \quad \text{oraz} \quad \eta = \frac{W_F-W_c}{kT}$$

Wielkość x jest zmienną pomocniczą. W ostatecznych wzorach będzie występowała wielkość η, która zgodnie z definicją określa położenie poziomu Fermiego w stosunku do dna pasma przewodnictwa wyrażone w jednostkach kT. Wprowadzając x i η, uzyskuje się następującą postać wyrażenia (8/8):

$$n_0 = 4\pi\left(\frac{2m_c kT}{h^2}\right)^{3/2} \int\limits_{0}^{\infty} \frac{x^{1/2}\,\mathrm{d}x}{1+\exp(x-\eta)} \tag{8/9}$$

co można również przedstawić następująco:

$$n_0 = N_c F_{1/2}(\eta) \tag{8/10}$$

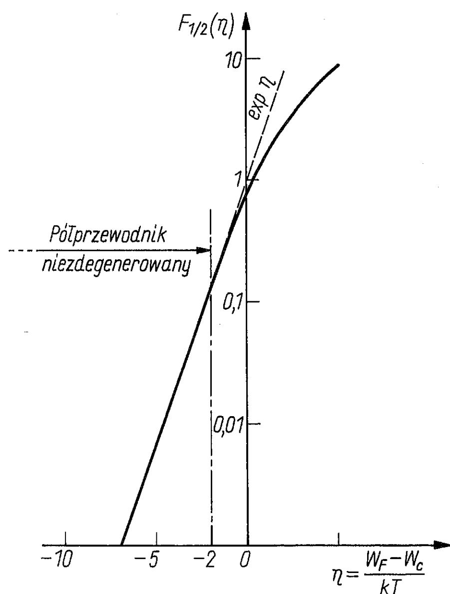
Wprowadzoną tu wielkość

$$N_c = 2\left(\frac{2\pi\, m_c\, kT}{h^2}\right)^{3/2} \tag{8/11}$$

nazywamy *efektywną gęstością stanów* w pasmie przewodnictwa, natomiast funkcja

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int\limits_0^\infty \frac{x^{1/2}\,dx}{1+\exp(x-\eta)} \tag{8/12}$$

nosi nazwę *całki Fermiego* o wskaźniku 1/2. Całka ta nie może być wyrażona przez funkcje elementarne. W postaci stabelaryzowanej jest ona podawana w niektórych monografiach[1]. Jej wykres pokazano na rys. 8/5.



8/5
Całka Fermiego $F_{1/2}(\eta)$

Podstawiając do wzoru (8/11) wartości stałych fizycznych, otrzymuje się następujący wzór liczbowy na efektywną gęstość stanów wyrażoną w cm^{-3}:

$$N_c = 2{,}5\cdot 10^{19}\left(\frac{m_c}{m_0}\right)^{3/2}\left(\frac{T}{300}\right)^{3/2} \tag{8/13}$$

Jak wynika z rys. 8/5, przy dostatecznie dużych ujemnych wartościach zmiennej η można funkcję $F_{1/2}(\eta)$ aproksymować funkcją wykładniczą, tzn. przyjmować, że

$$F_{1/2}(\eta) \approx \exp\eta \tag{8/14}$$

co jest równoznaczne z zastosowaniem przybliżenia boltzmannowskiego, czyli zastąpienia funkcji Fermiego-Diraca funkcją (8/3). Przy $\eta < -2$ niedokładność wynikająca z tego uproszczenia jest mniejsza niż 5%.

[1] Np. J. S. Blakemore: *Semiconductor statistics*. Pergamon Press, Oxford 1962, s. 351–353.

Półprzewodnik, dla którego spełniony jest warunek umożliwiający zastosowanie przybliżenia (8/14), nazywamy *niezdegenerowanym*[1]. Warunek $\eta < -2$ jest równoważny warunkowi

$$W_F < W_c - 2kT \tag{8/15}$$

Dla półprzewodnika niezdegenerowanego poziom Fermiego musi więc znajdować się w przerwie energetycznej w dostatecznie dużej odległości od dna pasma przewodnictwa (i, jak dalej zobaczymy, również w dostatecznie dużej odległości od wierzchołka pasma walencyjnego). Dla takiego półprzewodnika

$$\boxed{n_0 = N_c \exp\left(-\frac{W_c - W_F}{kT}\right)} \tag{8/16}$$

Koncentracja równowagowa elektronów w pasmie przewodnictwa zależy więc od rodzaju półprzewodnika, temperatury i położenia poziomu Fermiego w stosunku do dna tego pasma. Wzór (8/16) nie wystarcza jednak do obliczenia n_0, ponieważ położenie poziomu Fermiego nie zostało jeszcze wyznaczone. Jak zobaczymy w p. 8.6, jest ono zależne od rodzaju i koncentracji domieszek w półprzewodniku.

Problem 8.2. Jeżeli jest spełniony warunek (8/15), to dla energii zawartych w pasmie przewodnictwa można z niewielkim błędem stosować przybliżenie boltzmannowskie. Wykaż, że postępując tak, otrzymuje się ze wzoru (8/5) wzór (8/16).

Przechodząc z kolei do wyznaczenia koncentracji dziur w pasmie walencyjnym, skorzystamy z tego, że liczba dziur jest równa liczbie nie obsadzonych stanów w tym pasmie. Prawdopodobieństwo obsadzenia stanu o energii W przez dziurę wynosi więc $1-f(W)$, a koncentracja równowagowa dziur wyraża się wzorem

$$p_0 = \int_{-\infty}^{W_v} [1-f(W)]\, g(W)\, \mathrm{d}W \tag{8/17}$$

gdzie $g(W)$ jest gęstością stanów w pasmie walencyjnym. Przez analogię do wzoru (8/7) można napisać

$$g(W) = 4\pi \left(\frac{2m_v}{h^2}\right)^{3/2} (W_v - W)^{1/2} \tag{8/18}$$

przy czym m_v oznacza masę efektywną gęstości stanów dla tego pasma.

---

[1] Chodzi tu o inny rodzaj degeneracji niż w rozdz. 3, gdzie była mowa o zdegenerowanych poziomach energetycznych. Półprzewodnik niezdegenerowany jest to taki półprzewodnik, do którego zamiast statystyki Fermiego-Diraca można stosować statystykę Maxwella-Boltzmanna.

Opierając się na przekształceniu (8/2), dochodzimy do wniosku, że koncentracja równowagowa dziur w pasmie walencyjnym może być wyrażona wzorem

$$p_0 = N_v F_{1/2}\left(-\frac{W_F - W_v}{kT}\right) \tag{8/19}$$

gdzie

$$N_v = 2\left(\frac{2\pi m_v kT}{h^2}\right)^{3/2} \tag{8/20}$$

oznacza efektywną gęstość stanów w tym pasmie.

Dla półprzewodnika niezdegenerowanego, tj. takiego, dla którego jest spełniony warunek

$$W_F > W_v + 2kT \tag{8/21}$$

wzór (8/19) upraszcza się do postaci

$$\boxed{p_0 = N_v \exp\left(-\frac{W_F - W_v}{kT}\right)} \tag{8/22}$$

Podobnie jak wyrażona wzorem (8/16) koncentracja elektronów, tak i koncentracja równowagowa dziur zależy od rodzaju półprzewodnika, temperatury i położenia poziomu Fermiego.

Dla większości półprzewodników przerwa energetyczna jest większa niż 4 kT. W związku z tym nawet w przypadku degeneracji półprzewodnika jeden z warunków (8/15) lub (8/21) jest spełniony. W takim przypadku przy wyznaczaniu koncentracji nośników większościowych konieczne jest stosowanie statystyki Fermiego-Diraca, natomiast przy wyznaczaniu koncentracji nośników mniejszościowych można posługiwać się przybliżeniem boltzmannowskim, na którym oparte są wzory (8/16) i (8/22).

8.4 PRAWO DZIAŁANIA MAS

Rozważania zawarte w tym i w dwóch następnych punktach będą ograniczone do półprzewodników niezdegenerowanych, bądą więc oparte na wyprowadzonych w poprzednim punkcie wzorach (8/16) i (8/22).

Mnożąc te wzory przez siebie, otrzymujemy

$$n_0 p_0 = N_c N_v \exp\left(-\frac{W_g}{kT}\right) \tag{8/23}$$

gdzie $W_g = W_c - W_v$ jest przerwą energetyczną półprzewodnika.

We wzorze tym nie występuje energia W_F, a zatem dla półprzewodnika niezdegenerowanego w stanie równowagi termodynamicznej iloczyn $n_0 p_0$ zależy tylko od rodzaju materiału podstawowego (m_c, m_v, W_g) i od temperatury. Nie zależy natomiast od rodzaju i koncentracji domieszek.

Ten ważny wniosek jest określany nazwą *prawa działania mas*[1]. Z prawa tego wynika, że jeżeli np. wskutek wprowadzenia odpowiednich domieszek ulegnie zwiększeniu koncentracja elektronów, to w tym samym stopniu zmniejszy się koncentracja dziur i odwrotnie. Wzrost koncentracji nośników jednego rodzaju zwiększa bowiem intensywność procesu rekombinacji par elektron-dziura, a przez to zmniejsza koncentrację nośników drugiego rodzaju.

W półprzewodniku samoistnym elektrony i dziury są wytwarzane i zanikają parami. Ich koncentracje są więc jednakowe i są oznaczane symbolem n_i[2]. Wielkość tę nazywamy *koncentracją samoistną*.

Kładąc we wzorze (8/23) $n_0 = p_0 = n_i$, otrzymamy

$$n_i^2 = N_c N_v \exp\left(-\frac{W_g}{kT}\right) \tag{8/24}$$

Ponieważ prawe strony wzorów (8/23) i (8/24) są identyczne, zatem prawo działania mas może być zapisane w postaci

$$\boxed{n_0 p_0 = n_i^2} \tag{8/25}$$

8.5 PÓŁPRZEWODNIK SAMOISTNY

Ze wzoru (8/24) otrzymuje się następujące wyrażenie na koncentrację samoistną:

$$\boxed{n_i = \sqrt{N_c N_v} \exp\left(-\frac{W_g}{2kT}\right)} \tag{8/26}$$

Opierając się na wzorze (8/13) i analogicznym wyrażeniu określającym N_v, otrzymujemy następujący wzór liczbowy na koncentrację samoistną wyrażoną w cm^{-3}:

$$n_i = 2{,}5\cdot 10^{19}\left(\frac{m_c m_v}{m_0^2}\right)^{3/4}\left(\frac{T}{300}\right)^{3/2}\exp\left(-\frac{W_g}{2kT}\right) \tag{8/27}$$

We wzorze tym największe znaczenie ma czynnik wykładniczy $\exp(-W_g/2kT)$. Koncentracja samoistna zależy więc w bardzo silnym stopniu od wartości przerwy energetycznej i od temperatury. Przerwa energetyczna określa energię, jaką należy dostarczyć elektronowi w celu wzbudzenia go z pasma walencyjnego do pasma przewodnictwa. Im większa wartość W_g, tym mniejsza liczba wzbudzonych elektronów, a więc również mniejsza koncentracja par elektron-dziura.

---

[1] Nazwa tego prawa nawiązuje do analogicznego prawa chemicznego (zwanego również prawem równowagi chemicznej lub prawem Guldberga i Waagego), określającego ilościowo stan równowagi chemicznej, osiągany w stałej temperaturze przez reakcję odwracalną.

[2] Indeks „i" pochodzi od angielskiego słowa „intrinsic" (wewnętrzny, specyficzny).

Przykładowo, dla krzemu ($m_c = 1{,}18\, m_0$; $m_v = 0{,}81\, m_0$; $W_g = 1{,}12$ eV) w temperaturze pokojowej ($T = 300$ K) koncentracja ta wynosi

$$\boxed{n_i = 9{,}45 \cdot 10^9 \approx 10^{10}\ \mathrm{cm}^{-3}}$$

Ponieważ koncentracja atomów w krysztale krzemu wynosi $5 \cdot 10^{22}$ cm^{-3}, zatem 1 para elektron-dziura przypada na $5 \cdot 10^{12}$ (tzn. na 5 bilionów) atomów[1].

Dla germanu, który ma mniejszą przerwę energetyczną ($W_g = 0{,}66$ eV), koncentracja samoistna w temperaturze pokojowej wynosi

$$n_i = 2 \cdot 10^{13}\ \mathrm{cm}^{-3}$$

jest więc 2 tysiące razy większa niż dla krzemu.

Dla półprzewodników o dużej szerokości przerwy energetycznej koncentracja samoistna w temperaturze pokojowej jest znikomo mała. Przykładowo idealny kryształ diamentu ($W_g = 5{,}5$ eV), w którym występowałaby średnio jedna para elektron-dziura musiałby mieć wielkość kuli ziemskiej.

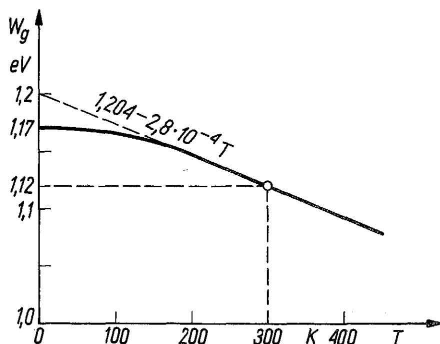
Z kolei powiększanie temperatury, będącej miarą średniej energii kinetycznej cząstek rozpatrywanego układu, powoduje wzrost koncentracji samoistnej. Im wyższa jest bowiem temperatura półprzewodnika, tym większa jest energia drgań atomów tworzących sieć krystaliczną, a wskutek tego większe prawdopodobieństwo zerwania wiązania międzyatomowego. Ze wzrostem temperatury wzrasta więc szybkość generacji par elektron-dziura, a to pociąga za sobą wzrost n_i.

Temperatura występuje we wzorze (8/27) w postaci jawnej, a ponadto masy efektywne m_c i m_v oraz przerwa energetyczna W_g zależą w pewnym stopniu od temperatury. Wpływ zmian mas efektywnych jest we wzorze (8/27) mało istotny, natomiast zmiany W_g muszą być uwzględnione, ponieważ, jak to już zaznaczono, koncentracja samoistna półprzewodnika jest bardzo szybko malejącą funkcją W_g. Przykład zależności przerwy energetycznej od temperatury podano na rys. 8/6.

Jak widać, przerwa energetyczna krzemu, podobnie jak większości innych półprzewodników, maleje ze wzrostem temperatury. W zakresie temperatur 200–500 K wartość W_g w eV dla krzemu może być wyrażona wzorem

$$W_g = 1{,}204 - 2{,}8 \cdot 10^{-4}\, \mathrm{T} \tag{8/28}$$

[1] Operując takimi liczbami, warto uświadomić sobie, że średnie życie ludzkie trwa zaledwie 2 miliardy sekund, a czas, który upłynął od początku wszechświata, jest oceniany na $4 \cdot 10^{17}$ s.



8/6
Zależność szerokości przerwy energetycznej krzemu od temperatury

Problem 8.3. Wyznacz zależność koncentracji samoistnej krzemu od temperatury w zakresie od 200 do 500 K.

Rozwiązanie: Wprowadzając (8/28) do wyrażenia $\exp(-W_g/2kT)$, otrzymuje się

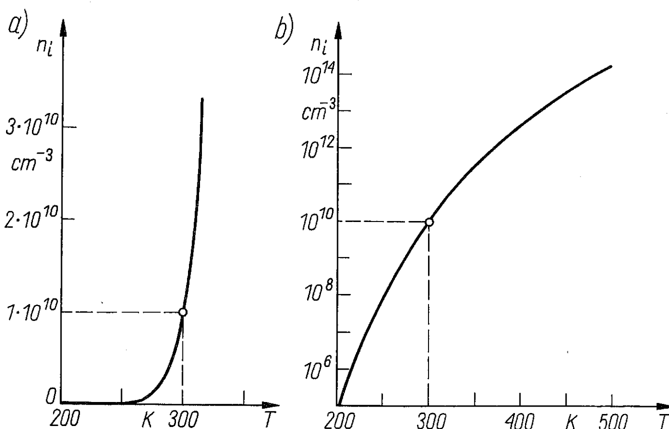
$$\exp\left(-\frac{W_g}{2kT}\right) = \exp\left(-\frac{1{,}204}{2kT}\right)\exp\left(\frac{2{,}8\cdot 10^{-4}}{2k}\right) \tag{8/29}$$

Stałą Boltzmanna k należy tu wyrazić w eV/K. Podstawiając do wzoru (8/27) odpowiednie wartości mas efektywnych, otrzymujemy po prostych przekształceniach

$$n_i = 2{,}36\cdot 10^{16}\, T^{3/2}\exp\left(-\frac{6986}{T}\right) \tag{8/30}$$

We wzorze tym wpływ czynnika $T^{3/2}$ jest niewielki w porównaniu z czynnikiem wykładniczym. Przykładowo, powiększenie temperatury z 300 do 400 K powoduje, że czynnik $T^{3/2}$ wzrasta o nieco ponad 50%, podczas gdy czynnik wykładniczy wzrasta w tych warunkach przeszło 300-krotnie.

Wykres zależności opisanej wzorem (8/30) przedstawiono na rys. 8/7. Dla porównania zestawiono tu wykresy sporządzone przy zastosowaniu na osi rzędnych skali liniowej i skali logarytmicznej. Ze względu na to, że w założonym przedziale temperatur koncentracja samoistna zmienia się o wiele rzędów wielkości, dogodniejszy jest wykres w skali logarytmicznej. Na wykresie w skali liniowej możliwe jest zaprezentowanie tylko niewielkiego odcinka rozpatrywanej charakterystyki.



8/7 Zależność koncentracji samoistnej krzemu od temperatury: a) skala liniowa; b) skala logarytmiczna

Na podstawie wyników pomiarów uzyskanych przez różnych autorów został opracowany następujący wzór na koncentrację samoistną krzemu (wyrażoną w cm^{-3}) w zakresie temperatur 200–500 K[1]:

$$n_i = 5{,}71\cdot 10^{19}\left(\frac{T}{300}\right)^{2{,}365}\exp\left(-\frac{6733}{T}\right) \tag{8/31}$$

Wartości obliczone na podstawie wzorów (8/30) i (8/31) pokrywają się z dokładnością do kilku procent przy temperaturach nie odbiegających zbytnio od temperatury pokojowej.

[1] Th. Wasserrab: Die Temperaturabhängigkeit der elektronischen Kenngrößen des eigenleiten den Siliziums. *Z. Naturfosch.* **32a**, 746 (1977).

Problem 8.4. Wyprowadź ogólny wzór na współczynnik temperaturowy koncentracji samoistnej. Współczynnik ten, będący *pochodną logarytmiczną* funkcji $n_i(T)$, zdefiniowany jako

$$\frac{1}{n_i}\frac{\mathrm{d}n_i}{\mathrm{d}T} = \frac{\mathrm{d}}{\mathrm{d}T}(\ln n_i)$$

określa względną zmianę n_i wywołaną zmianą temperatury o 1 K. Załóż, że przerwa energetyczna półprzewodnika zmienia się z temperaturą wg wzoru

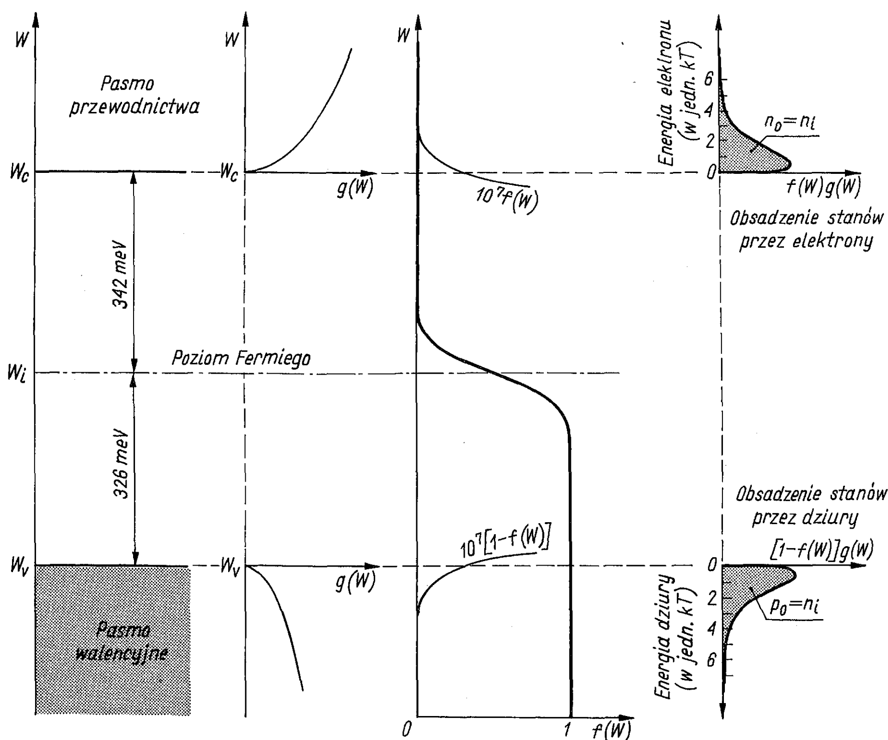
$$W_g(T) = W_{g0} + \alpha T$$

gdzie W_{g0} – ekstrapolowana do $T = 0$ wartości W_g (zob. rys. 8/6). Wyznacz wartość tego współczynnika dla krzemu w temperaturze pokojowej.

Na podstawie otrzymanych uprzednio wzorów można również wyznaczyć położenie poziomu Fermiego w półprzewodniku samoistnym. Poziom ten nazywamy *poziomem samoistnym* i oznaczamy symbolem W_i. Z warunku $n_0 = p_0$, a więc z przyrównania prawych stron wzorów (8/16) i (8/22) otrzymuje się

$$W_i = \frac{W_c + W_v}{2} - \frac{1}{2}kT\ln\frac{N_c}{N_v} = \frac{W_c + W_v}{2} - \frac{3}{4}kT\ln\frac{m_c}{m_v} \qquad (8/32)$$

W półprzewodniku samoistnym poziom Fermiego znajduje się więc niemal dokładnie w środku przerwy energetycznej. Niewielkie przesunięcie poziomu



8/8 Przebieg funkcji g (W) i *f* (W) oraz obsadzenia stanów w pasmie przewodnictwa i w pasmie walencyjnym germanu samoistnego w temperaturze $T = 300$ K

W_i w stosunku do środka przerwy jest spowodowane różnicą mas efektywnych m_c i m_v. W temperaturze pokojowej wynosi ono dla germanu 8 meV, a dla krzemu 7 meV.

Znajomość położenia poziomu Fermiego umożliwia sporządzenie podanych na rys. 8/8 wykresów przedstawiających rozkład gęstości stanów w pasmie przewodnictwa i w pasmie walencyjnym oraz przebieg funkcji rozkładu i obsadzenia stanów w obu tych pasmach. Przebieg funkcji rozkładu $f(W)$ przedstawiono w bardzo szerokim zakresie energii. Z tego względu na wykresie tym nie można odczytać wartości $f(W)$ dla $W \geqslant W_c$, ani też wartości $1-f(W)$ dla $W \leqslant W_v$. Aby to umożliwić podano tam dodatkowe wykresy obu tych funkcji w skali powiększonej 10^7 razy.

8.6 PÓŁPRZEWODNIKI DOMIESZKOWE

Z rozważań zawartych w p. 8.3, a w szczególności ze wzorów (8/16) i (8/22) wynika, że o koncentracjach nośników w danym półprzewodniku i w danej temperaturze decyduje położenie poziomu Fermiego. W półprzewodniku samoistnym koncentracje elektronów i dziur są jednakowe. Z warunku tego wyciągnęliśmy wniosek dotyczący położenia poziomu samoistnego W_i. Dla takiego półprzewodnika wymienione wzory przyjmują postać:

$$n_0 = N_c \exp\left(-\frac{W_c - W_i}{kT}\right) = n_i \tag{8/33}$$

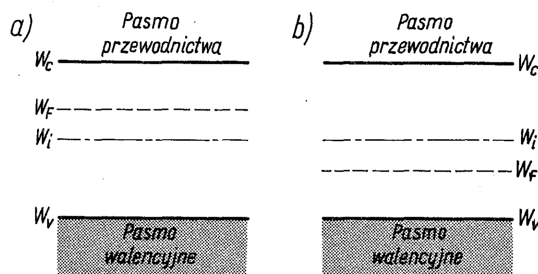
$$p_0 = N_v \exp\left(-\frac{W_i - W_v}{kT}\right) = n_i \tag{8/34}$$

Korzystając z nich, można wyrażenia (8/16) i (8/22) doprowadzić do postaci

$$n_0 = n_i \exp\frac{W_F - W_i}{kT} \tag{8/35}$$

$$p_0 = n_i \exp\frac{W_i - W_F}{kT} \tag{8/36}$$

Z pierwszego z tych wzorów wynika, że jeżeli w rozpatrywanym półprzewodniku poziom Fermiego znajduje się powyżej poziomu samoistnego (rys. 8/9a), to koncentracja elektronów jest większa niż koncentracja samoistna n_i, i to tym bardziej, im większa jest różnica $W_F - W_i$, a więc im wyżej w modelu pasmowym znajduje się poziom W_F. Jednocześnie zgodnie ze wzorem (8/36) i z prawem działania mas, koncentracja dziur jest tyle samo razy mniejsza od n_i. Ponieważ w temperaturze pokojowej $kT = 26$ meV, zatem każde przesunięcie poziomu W_F w górę o 60 meV wywołuje 10-krotne zwiększenie n_0 i 10-krotne zmniejszenie p_0.



8/9 Modele pasmowe półprzewodników domieszkowych: a) półprzewodnik typu *n*; b) półprzewodnik typu *p*

Jak już powiedziano w p. 7.2, półprzewodnik, w którym koncentracja elektronów jest większa od koncentracji dziur, nazywamy *półprzewodnikiem typu n*. W półprzewodniku takim elektrony są nazywane nośnikami *większościowymi*, dziury zaś – nośnikami *mniejszościowymi*.

Analogicznie, jeżeli poziom W_F znajduje się poniżej poziomu samoistnego (rys. 8/9b), to koncentracja dziur jest większa od koncentracji samoistnej, a koncentracja elektronów tyle samo razy mniejsza od n_i. Półprzewodnik taki nazywamy *półprzewodnikiem typu p*. Nośnikami większościowymi są w nim dziury, a nośnikami mniejszościowymi – elektrony.

Istnieją 2 sposoby wpływania na położenie poziomu Fermiego w półprzewodniku. Jednym z nich jest domieszkowanie go atomami, które łatwo ulegają jonizacji w sieci krystalicznej, drugim zaś oddziaływanie na półprzewodnik zewnętrznym polem elektrycznym.

W książce tej nie będziemy się zajmowali zjawiskami związanymi z oddziaływaniem zewnętrznego pola elektrycznego na półprzewodnik, rozpatrzymy tu natomiast wpływ domieszek na koncentracje równowagowe nośników i na położenie poziomu Fermiego w półprzewodniku. W tym celu musimy wziąć pod uwagę koncentracje poszczególnych rodzajów ładunków elektrycznych, z jakimi mamy do czynienia w półprzewodniku w stanie równowagi termodynamicznej.

W ogólnym przypadku występują w półprzewodniku 4 rodzaje ładunków elektrycznych:

a) zjonizowane domieszki: donory i akceptory;

b) nośniki ładunku, tj. swobodne elektrony i dziury.

Ładunki te powstają w półprzewodniku zawsze parami – np. wzbudzeniu elektronu do pasma przewodnictwa towarzyszy albo pojawienie się dziury, albo też zjonizowanego donoru. Z tego względu w półprzewodniku jednorodnym w warunkach równowagi termodynamicznej jest spełnione następujące równanie orzekające, że koncentracje ładunków dodatnich i ujemnych są sobie równe:

$$p_0+N_d^+ = n_0+N_a^- \tag{8/37}$$

Równanie to wyraża tzw. *warunek neutralności elektrycznej* półprzewodnika.

Nasuwa się tu pytanie, jaka część wszystkich donorów i akceptorów jest

zjonizowana. Stopień jonizacji tych domieszek zależy przede wszystkim od temperatury, a w pewnym stopniu i w pewnym zakresie również od ich koncentracji[1]. Jeżeli jednak temperatura nie odbiega zbytnio od temperatury pokojowej (tzn. jeśli jest spełniony warunek $T > 250$ K), to z niewielkim błędem można przyjmować, że niezależnie od koncentracji wszystkie domieszki są zjonizowane, tzn.

$$N_d^+ = N_d \quad \text{oraz} \quad N_a^- = N_a$$

Dla krzemu maksymalna niedokładność tych równości występuje przy koncentracjach donorów lub akceptorów leżących w zakresie od 10^{18} do 10^{19} cm^{-3}. Niedokładność tę można oceniać na kilka do kilkunastu procent. Błąd wynikający z założenia całkowitej jonizacji domieszek nie jest więc zbyt duży, zwłaszcza jeśli uwzględnimy fakt, że dokładność, z jaką jest znana koncentracja domieszek w półprzewodnikach wynosi w najlepszym przypadku 5–10%.

Dalsze rozważania zawarte w tym punkcie będą ograniczone do zakresu temperatur wyższych niż 250 K, będą zatem oparte na założeniu całkowitej jonizacji domieszek.

Z przeprowadzonych rozważań wynika, że dysponujemy dwoma równaniami, w których występują szukane wielkości n_0 i p_0. Są to: warunek neutralności elektrycznej – wzór (8/37) oraz prawo działania mas – wzór (8/25). Rozwiązanie tych równań przy założeniu całkowitej jonizacji domieszek prowadzi do następujących wzorów:

$$n_0 = \frac{N_d - N_a}{2} + \sqrt{\left(\frac{N_d - N_a}{2}\right)^2 + n_i^2} \tag{8/38}$$

$$p_0 = \frac{N_a - N_d}{2} + \sqrt{\left(\frac{N_a - N_d}{2}\right)^2 + n_i^2} \tag{8/39}$$

O koncentracji nośników w półprzewodniku domieszkowym w stanie równowagi termodynamicznej decyduje więc różnica $N_d - N_a$. Wielkość tę będziemy oznaczali symbolem N i nazywali *efektywną koncentracją domieszek*. Zgodnie z definicją wynosi ona

$$N = N_d - N_a \tag{8/40}$$

Jeżeli $N_d > N_a$, to $N > 0$, i w wyniku tego $n_0 > n_i$. W tym przypadku mamy więc do czynienia z półprzewodnikiem typu n. Koncentracja elektronów w takim półprzewodniku zależy od nadwyżki koncentracji donorów nad koncentracją akceptorów. Elektrony uwolnione z poziomów donorowych obsadzają bowiem przede wszystkim poziomy akceptorowe, a po zapełnieniu wszystkich tych poziomów obsadzają stany w pasmie przewodnictwa półprzewodnika.

Jeżeli $N > 3\,n_i$, to z niewielkim błędem (< 1%) można wzór (8/38) uprościć do postaci

$$n_0 = N + \frac{n_i^2}{N} \tag{8/41}$$

Koncentrację dziur można następnie wyznaczyć, korzystając z prawa działania mas.

---

[1] Bardziej szczegółowo zagadnienie to jest omówione w dodatku E.

W półprzewodniku domieszkowym koncentracja domieszek jednego rodzaju przewyższa zazwyczaj o kilka rzędów wielkości koncentrację samoistną i koncentrację domieszek drugiego rodzaju. Jeżeli więc dla półprzewodnika typu n są spełnione warunki $N_d \gg n_i$ oraz $N_d \gg N_a$, to koncentracje równowagowe wynoszą odpowiednio

$$n_0 = N_d \tag{8/42}$$

$$p_0 = \frac{n_i^2}{N_d} \tag{8/43}$$

Podobnie w przypadku $N_a > N_d$, otrzymujemy $N < 0$ i co za tym idzie $p_0 > n_i$. Rozpatrywany materiał jest więc półprzewodnikiem typu p, a o koncentracjach nośników decyduje nadwyżka koncentracji akceptorów nad koncentracją donorów. Wypływa stąd wniosek, że półprzewodnik typu n można przekształcić w półprzewodnik typu p (bez zmiany ilości zawartych w nim donorów), wprowadzając do niego odpowiednią ilość akceptorów i odwrotnie. Metoda ta jest powszechnie stosowana w technologii elementów półprzewodnikowych.

W półprzewodniku typu p, w którym $|N| > 3\,n_i$

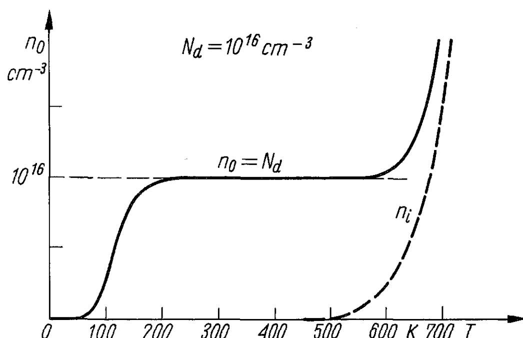
$$p_0 = |N| + \frac{n_i^2}{|N|} \tag{8/44}$$

Jeżeli natomiast są spełnione warunki $N_a \gg n_i$ oraz $N_a \gg N_d$, to

$$p_0 = N_a \tag{8/45}$$

$$n_0 = \frac{n_i^2}{N_a} \tag{8/46}$$

W półprzewodniku domieszkowym, w którym dominuje jeden rodzaj domieszek, przy czym ich koncentracja przewyższa znacznie n_i, koncentracja nośników większościowych jest zatem równa koncentracji domieszek. Jest ona więc niezależna od rodzaju półprzewodnika i od temperatury, w takim oczywiście zakresie, w którym są spełnione założenia, na których opierają się wzory (8/42) i (8/45). Pokazano to na rys. 8/10, gdzie przedstawiono koncentrację elektronów



8/10
Przykład zależności koncentracji równowagowej elektronów w krzemie typu *n* od temperatury

w półprzewodniku typu n w zależności od temperatury. Jak widać, w szerokim zakresie temperatur $n_0 = N_d$. Przy bardzo niskich temperaturach koncentracja elektronów jest mniejsza niż koncentracja donorów, ponieważ nie wszystkie donory są w tych warunkach zjonizowane. Z kolei przy dostatecznie wysokich temperaturach koncentracja samoistna staje się porównywalna z koncentracją donorów, w związku z czym warunek $N_d \gg n_i$, na którym opiera się wzór (8/42), przestaje być spełniony.

Koncentracja nośników mniejszościowych zależy natomiast w bardzo silnym stopniu od rodzaju półprzewodnika i od temperatury. Jest ona bowiem, jak to wynika ze wzorów (8/43) i (8/46) proporcjonalna do kwadratu koncentracji samoistnej, a ta, jak wiemy, jest bardzo wrażliwa na zmiany wartości przerwy energetycznej półprzewodnika i temperatury.

Jeżeli przewaga koncentracji donorów nad koncentracją akceptorów, lub odwrotnie, jest bardzo mała, tzn. jeżeli $|N_d - N_a| \ll n_i$, to, jak wynika ze wzorów (8/38) i (8/39)

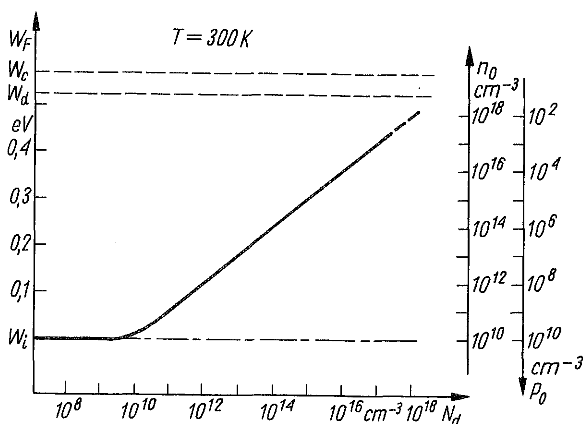
$$n_0 \approx p_0 \approx n_i$$

Elektrony z poziomów donorowych obsadzają wówczas wszystkie lub prawie wszystkie poziomy akceptorowe, w wyniku czego koncentracje elektronów w pasmie przewodnictwa i dziur w pasmie walencyjnym są w przybliżeniu takie, jak w półprzewodniku samoistnym. Półprzewodnik zawierający w przybliżeniu jednakową ilość donorów i akceptorów nazywamy półprzewodnikiem *skompensowanym.*

Pozostało nam jeszcze do wyznaczenia położenie poziomu Fermiego w półprzewodniku domieszkowym. Posłużymy się tu wyrażeniem (8/35), które można przedstawić w postaci

$$W_F - W_i = kT \ln \frac{n_0}{n_i} \tag{8/47}$$

Wzór ten jest słuszny zarówno dla półprzewodników typu n, jak i typu p. Występującą w nim koncentrację n_0 należy w ogólnym przypadku wyznaczać ze wzoru (8/38). Przebieg zależności W_F od koncentracji donorów dla półprzewodnika typu n przedstawiono na rys. 8/11. Pokazano tam również jak zmieniają



8/11 Zależność położenia poziomu Fermiego oraz równowagowych koncentracji nośników w krzemie typu n od koncentracji donorów

się koncentracje równowagowe n_0 i p_0 w funkcji N_d. Wszystkie 3 zależności, dzięki odpowiedniemu doborowi skal na osiach rzędnych, przedstawiono za pomocą jednej krzywej.

Wykres pokazany na rys. 8/11 dotyczy krzemu w temperaturze pokojowej. Wynika z niego, że przy koncentracjach $N_d \leqslant 10^9\ \mathrm{cm}^{-3}$ półprzewodnik ten można uważać za samoistny. Poziom Fermiego pokrywa się wówczas, praktycznie rzecz biorąc, z poziomem samoistnym, a koncentracje nośników są w przybliżeniu równe n_i[1]. Po przekroczeniu wartości $N_d = 10^9\ \mathrm{cm}^{-3}$ poziom W_F przesuwa się w kierunku pasma przewodnictwa. Dla $N_d \geqslant 10^{11}\ \mathrm{cm}^{-3}$ można z błędem mniejszym niż 1% przyjmować, że $n_0 = N_d$, wskutek czego wzór (8/47) sprowadza się do postaci

$$\boxed{W_F - W_i = kT \ln \frac{N_d}{n_i}} \qquad (8/48)$$

Przy dalszym powiększaniu koncentracji donorów poziom Fermiego zbliża się coraz bardziej do poziomów donorowych W_d. Zwiększa się więc prawdopodobieństwo obsadzenia tych poziomów przez elektrony. W wyniku tego z coraz gorszym przybliżeniem jest spełnione założenie całkowitej jonizacji domieszek, na którym opierają się rozważania zawarte w tym punkcie. Poziom Fermiego zbliża się także do dna pasma przewodnictwa i przy pewnej koncentracji donorów półprzewodnik staje się zdegenerowany. Istnieje więc pewna maksymalna wartość N_d, po przekroczeniu której wzory wyprowadzone w tym punkcie przestają być słuszne.

Problem 8.5. Dla krzemu domieszkowanego fosforem wyznacz koncentrację atomów fosforu, przy której w temperaturze pokojowej półprzewodnik ten staje się zdegenerowany.

Położenie poziomu Fermiego w półprzewodniku typu p może być, jak to już zaznaczono, wyznaczone ze wzoru (8/47). Jeżeli koncentracja akceptorów jest dostatecznie duża (tzn. $N_a \gg N_d$ oraz $N_a \gg n_i$), wówczas wygodniej jest posługiwać się następującym wzorem analogicznym do wyrażenia (8/48):

$$\boxed{W_i - W_F = kT \ln \frac{N_a}{n_i}} \qquad (8/49)$$

8.7* ZJAWISKA WYSTĘPUJĄCE PRZY DUŻYCH KONCENTRACJACH DOMIESZEK[2]

Rozważania zawarte w poprzednich punktach tego rozdziału odnoszą się do półprzewodników, w których koncentracje domieszek nie są zbyt duże. Po prze-

[1] Dokładne obliczenia prowadzą do wniosku, że dla $N_d = 10^9\ \mathrm{cm}^{-3}$, $n_0 = 1{,}05\, n_i$, a $p_0 = 0{,}95\, n_i$.

[2] Punkt ten jest w znacznym stopniu oparty na pracy: W. Kuźmicz: Zjawiska występujące w półprzewodnikach silnie domieszkowanych i ich wpływ na parametry przyrządów półprzewodnikowych. *Prace naukowe Politechniki Warszawskiej*. Elektronika. Z. 61 (1985).

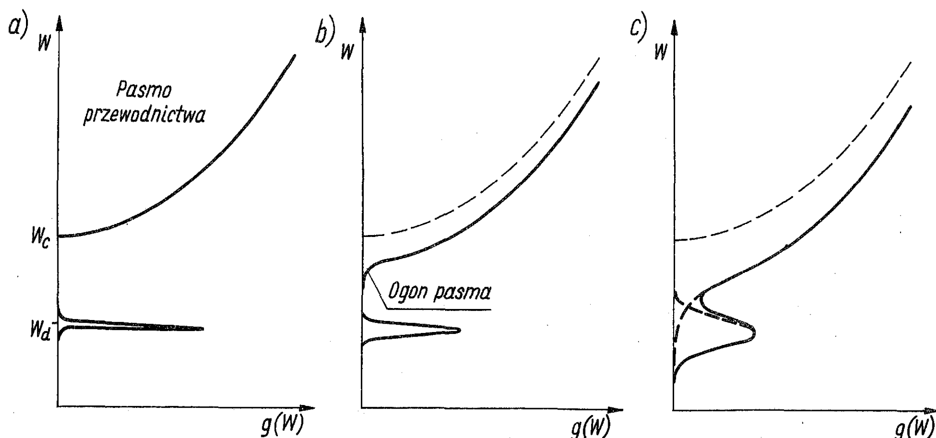
kroczeniu pewnej wartości koncentracji, która np. dla krzemu wynosi ok. 10^{17} cm^{-3}, zaczynają odgrywać rolę zjawiska, których dotychczas nie braliśmy pod uwagę. Koncentracje domieszek tego rzędu osiąga się, a nawet znacznie przekracza w niektórych obszarach większości elementów półprzewodnikowych. W punkcie tym zajmiemy się zatem zagadnieniem koncentracji równowagowych w silnie domieszkowanych półprzewodnikach. Rozważania nasze będą dotyczyły głównie krzemu, ze względu na to, że większość opublikowanych prac z tego zakresu dotyczy tego właśnie materiału. W celu ustalenia uwagi rozpatrzymy tu krzem typu *n*, w którym jedynymi domieszkami są donory.

Omawiane w rozdziałach 6 i 7 struktury energetyczne półprzewodników dotyczą kryształów idealnych o doskonale regularnej sieci. Obecność w sieci krystalicznej obcych atomów, mających inne rozmiary niż atomy własne, a także inny rozkład potencjału wokół rdzenia atomowego, zakłóca tę regularność i wskutek tego wywołuje zmianę struktury energetycznej półprzewodnika. Przy dostatecznie małych koncentracjach domieszek zmiany te są niedostrzegalne dzięki czemu gęstości stanów w pobliżu krawędzi pasm energetycznych mogą być wyrażone wzorami (8/7) i (8/18). W miarę wzrostu koncentracji domieszek wzory te stają się coraz mniej dokładne. W strukturze energetycznej półprzewodnika pojawiają się tzw. *„ogony” pasm*, tzn. dodatkowę stany kwantowe zlokalizowane bezpośrednio pod poziomem W_c i ponad poziomem W_v. Gęstości tych stanów maleją w przybliżeniu wykładniczo w miarę przesuwania się od pierwotnej krawędzi pasma w kierunku środka przerwy energetycznej. W tych warunkach ostre granice oddzielające pasma energii dozwolonych od przerwy energetycznej przestają istnieć, a pojęcie przerwy energetycznej traci swój jednoznaczny sens.

Innym efektem wywołanym przez wprowadzenie do półprzewodnika dodatkowych obcych atomów jest pojawienie się w krysztale naprężeń mechanicznych oraz różnego rodzaju defektów strukturalnych. Wywołują one dodatkowe zmiany struktury energetycznej półprzewodnika. Te skutki domieszkowania mogą być w znacznym stopniu ograniczone przez zastosowanie odpowiednich zabiegów technologicznych. Na przykład wygrzewanie kryształu zmniejsza w istotny sposób naprężenia mechaniczne, usuwa część defektów strukturalnych i sprawia, że atomy domieszkowane zajmują w sieci położenia węzłowe. Z tego względu wygrzewanie kryształów po wprowadzeniu domieszek jest powszechnie stosowanym procesem porządkującym sieć krystaliczną półprzewodnika.

Kolejnym zjawiskiem występującym w silnie domieszkowanych półprzewodnikach jest przekształcenie się dyskretnych poziomów domieszkowych w pasma domieszkowe, o czym była już mowa w p. 7.6.2. W półprzewodniku typu *n* przy małych koncentracjach donorów odległości między atomami domieszkowymi są duże, a zatem elektrony związane z centrami donorowymi prawie nie oddziałują na siebie. W tych warunkach podstawowy poziom donorowy jest poziomem dyskretnym o ściśle określonej energii, którą oznaczamy symbolem W_d – zob. rys. 8/12a.

W miarę wzrostu koncentracji donorów funkcje falowe stanów donorowych zaczynają w coraz większym stopniu zachodzić na siebie. W wyniku tego poziom



8/12 Funkcja gęstości stanów w pasmie przewodnictwa i w pasmie domieszkowym dla półprzewodnika typu *n* przy małym (a), średnim (b) i dużym (c) poziomie domieszkowania

donorowy przekształca się w pasmo, którego szerokość wzrasta w miarę zwiększania N_d. Jednocześnie, jak o tym mówiliśmy w p. 7.6.2, pasmo przewodnictwa przesuwa się w kierunku środka przerwy energetycznej – zob. rys. 8/12b. W wyniku rozszerzania się pasma domieszkowego przy jednoczesnym obniżaniu się dna pasma przewodnictwa zmniejsza się energia jonizacji donorów, jak to pokazano na rys. 7/22. Przy pewnej koncentracji donorów (wynoszącej np. dla krzemu ok. $2 \cdot 10^{18}$ cm^{-3}) energia ta spada do zera – pasmo domieszkowe łączy się więc z pasmem przewodnictwa, a przy dalszym wzroście N_d zostaje ono stopniowo wchłonięte przez to ostatnie – zob. rys. 8/12c.

Zastanówmy się teraz nad tym, jaki wpływ mają opisane zjawiska na koncentracje równowagowe nośników. Tak jak już zaznaczyliśmy, dyskusja nasza będzie dotyczyła krzemu typu *n*. Przy koncentracjach donorów nie przekraczających wartości 10^{17} cm^{-3}, w temperaturze nie odbiegającej zbytnio od temperatury pokojowej koncentracje nośników wyrażają się wzorami (8/42) i (8/43). W miarę wzrostu N_d poziom Fermiego przesuwa się, jak wiemy, w kierunku pasma przewodnictwa, a zatem prawdopodobieństwo obsadzenia przez elektrony poziomów donorowych stopniowo wzrasta. Stopień jonizacji tych domieszek maleje więc, a koncentracja n_0 staje się mniejsza od N_d. Z chwilą wytworzenia się pasma domieszkowego, elektrony, których energie są zawarte w tym pasmie, uzyskują jednakże pewną swobodę przemieszczania się w całej objętości półprzewodnika. Stopień tej swobody zwiększa się stopniowo w miarę tego, jak pasmo domieszkowe zbliża się do dna pasma przewodnictwa, następnie łączy się z nim i w końcu zostaje przez nie wchłonięte. Przy koncentracjach N_d większych niż 10^{19} cm^{-3} można przyjmować, że

$$n_0 = N_d \tag{8/50}$$

co formalnie rzecz biorąc, jest równoważne założeniu, że wszystkie domieszki są zjonizowane.

Równość (8/42) obowiązuje zatem dla krzemu w temperaturze bliskiej temperatury pokojowej przy koncentracjach donorów mniejszych niż 10^{17} cm^{-3} lub większych niż 10^{19} cm^{-3}. W zakresie N_d leżącym między tymi wartościami jest ona obarczona pewnym błędem, który, jak to zaznaczono w p. 8.6, wynosi od kilku do kilkunastu procent.

Wzór (8/43) może być natomiast stosowany jedynie przy dostatecznie małych koncentracjach donorów (np. przy $N_d \leqslant 10^{17}$ cm^{-3}). Opiera się on bowiem na prawie działania mas, które przy większych koncentracjach domieszek przestaje obowiązywać. Prawo to wywodzi się mianowicie ze wzoru (8/24), który dla silnie domieszkowanych półprzewodników przestaje być słuszny z następujących 2 powodów: po pierwsze, jak to wynika z rozważań przeprowadzonych na początku tego punktu, w miarę wzrostu koncentracji donorów pasmo przewodnictwa przesuwa się w kierunku pasma walencyjnego, a zatem szerokość przerwy energetycznej maleje, a wskutek tego iloczyn $n_0 p_0$ wzrasta. Po drugie zaś, wzór (8/24) opiera się na założeniu, że półprzewodnik jest niezdegenerowany, a więc obowiązuje w nim statystyka Maxwella-Boltzmanna. Przy dużych koncentracjach domieszek założenie to nie jest, jak wiemy, spełnione – zob. probl. 8.5.

Zjawisko zmiany szerokości przerwy energetycznej można opisać wzorem

$$W_g = W_{g0} - \Delta W_g \tag{8/51}$$

gdzie: W_{g0} – szerokość przerwy w półprzewodniku nie domieszkowanym; ΔW_g zaś – zwężenie przerwy spowodowane wprowadzeniem domieszek.

Wzór (8/51) ma charakter formalny. Wielkość ΔW_g odzwierciedla bowiem nie tylko rzeczywiste zwężenie przerwy energetycznej, lecz również wpływ innych czynników powodujących zmiany iloczynu $n_0 p_0$ w zależności od koncentracji domieszek. Jednym z takich czynników są np. zmiany funkcji gęstości stanów.

Wartości ΔW_g podawane przez różnych autorów wykazują znaczny rozrzut, wynikający m.in. z niejednoznaczności pojęcia przerwy energetycznej w silnie domieszkowanym półprzewodniku, jak również z faktu, że bezpośredni pomiar wartości W_g, decydującej o iloczynie $n_0 p_0$, jest niemożliwy. Zestawienie wartości ΔW_g dla krzemu w zależności od koncentracji donorów przedstawiono na rys. 8/13. Krzywa pokazana na tym rysunku jest opisana wzorem

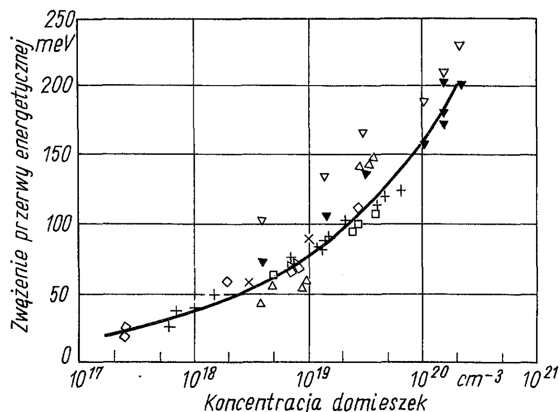
$$\Delta W_g = \left(\frac{N_d}{1{,}22 \cdot 10^{13}}\right)^{0{,}319} \tag{8/52}$$

przy czym ΔW_g jest wyrażone w meV.

W celu wyznaczenia iloczynu $n_0 p_0$ dla silnie domieszkowanego półprzewodnika typu n oprzemy się na wzorach (8/10) i (8/22). W przypadku nośników mniejszościowych stosujemy tu przybliżenie boltzmannowskie, ponieważ poziom Fermiego znajduje się w dużej odległości od wierzchołka pasma walencyjnego.

Po podstawieniu do (8/22) wyrażenia $W_v = W_c - W_g$ i po pewnych przekształceniach, otrzymujemy

$$p_0 = N_v \exp\left(-\eta - \frac{W_g}{kT}\right) \tag{8/53}$$



8/13 Zwężenie przerwy energetycznej w krzemie w zależności od koncentracji domieszek (Zaczerpnięto z: W. Zagożdżon-Wosik, W. Kuźmicz: Physical mechanisms of bandgap — narrowing in silicon. *Electron. Lett.* **19**, 515 (1983)

Mnożąc stronami wyrażenia (8/10) i (8/53) i korzystając ze wzoru (8/51), otrzymamy

$$n_0 p_0 = N_c N_v \exp\left(-\frac{W_{g0}}{kT}\right) \exp\left(\frac{\Delta W_g}{kT}\right) \frac{F_{1/2}(\eta)}{\exp \eta} \qquad (8/54)$$

Iloczyn pierwszych 3 czynników po prawej stronie równania jest to kwadrat koncentracji samoistnej n_i. Dwa ostatnie czynniki są natomiast odpowiedzialne za zmiany wartości iloczynu $n_0 p_0$ w funkcji koncentracji domieszki donorowej. Pierwszy z nich reprezentuje efekty związane ze zjawiskiem zwężenia przerwy energetycznej, drugi zaś stanowi poprawkę uwzględniającą wpływ degeneracji półprzewodnika, a więc poprawkę związaną z koniecznością stosowania statystyki Fermiego-Diraca.

Wzór (8/54) można napisać w postaci

$$\boxed{n_0 p_0 = n_{ie}^2} \qquad (8/55)$$

przy czym wielkość n_{ie}, nosząca nazwę *efektywnej koncentracji samoistnej*, jest identyczna z n_i przy dostatecznie małych koncentracjach domieszek, różni się natomiast od tej ostatniej dla silnie domieszkowanego półprzewodnika.

Wzór pozwalający na wyznaczenie n_{ie} jest dogodnie przedstawić w postaci

$$n_{ie}^2 = n_i^2 \exp \frac{\Delta W_{ge}}{kT} \qquad (8/56)$$

gdzie ΔW_{ge} — *efektywne zwężenie przerwy energetycznej* zdefiniowane wzorem

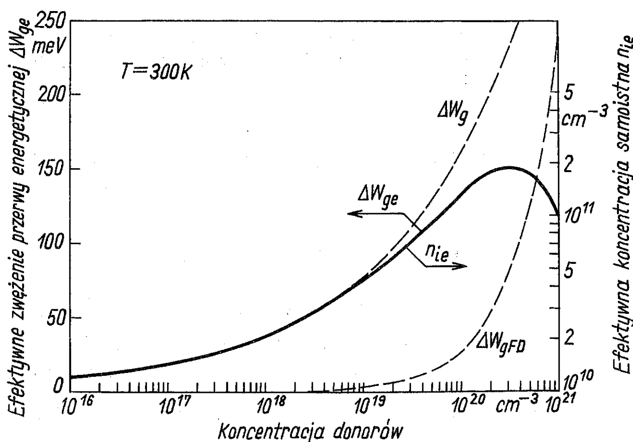
$$\Delta W_{ge} = \Delta W_g - \Delta W_{gFD} \qquad (8/57)$$

Wielkość ΔW_{gFD} reprezentuje pozorne rozszerzenie przerwy energetycznej, którego uwzględnienie pozwala zamiast statystyki Fermiego-Diraca stosować

statystykę Maxwella-Boltzmanna. Z przytoczonych poprzednio wzorów wynika, że wielkość ta wyraża się wzorem

$$\Delta W_{gFD} = kT \frac{\exp \eta}{F_{1/2}(\eta)} \tag{8/58}$$

Przebieg wszystkich 3 wielkości występujących we wzorze (8/57) w zależności od koncentracji donorów przedstawiono na rys. 8/14. Jak widzimy, efektywne zwężenie przerwy energetycznej ΔW_{ge} jest zawsze dodatnie, a zatem iloczyn $n_0 p_0$ jest dla silnie domieszkowanego półprzewodnika większy niż n_i^2. Dzięki zastosowaniu na rys. 8/14 odpowiednich skal na osi rzędnych, pokazano tam jednocześnie przebieg n_{ie} w funkcji N_d.



8/14 Efektywne zwężenie przerwy energetycznej i efektywna koncentracja samoistna dla krzemu typu *n* w zależności od koncentracji donorów

Pomimo formalnego charakteru wprowadzonych tu pojęć i zależności, przyjęło się w literaturze, że zamiast zmian wartości iloczynu $n_0 p_0$ w zależności od koncentracji domieszek są podawane równoważne zmiany ΔW_g, a zjawisko zmiany $n_0 p_0$ jest interpretowane jako efekt wywołany zwężeniem przerwy energetycznej. Jedną z głównych przyczyn takiego postępowania jest to, że iloczyn $n_0 p_0$ jest w bardzo silnym stopniu zależny od temperatury, podczas gdy wielkości ΔW_g i ΔW_{gFD} są prawie od niej niezależne.

9

GENERACJA I REKOMBINACJA NOŚNIKÓW NADMIAROWYCH W PÓŁPRZEWODNIKU

W poprzednim rozdziale zajmowaliśmy się koncentracjami nośników ładunku w półprzewodniku w stanie równowagi termodynamicznej. Stan ten może być traktowany jako punkt odniesienia do dalszych rozważań dotyczących *zjawisk nierównowagowych.* Są to warunki znacznie bardziej interesujące, ponieważ każde zjawisko fizyczne, które można obserwować lub mierzyć, jest związane z odstępstwem układu do równowagi termodynamicznej. Także działanie wszelkich elementów półprzewodnikowych, takich np. jak tranzystory, fotodetektory, diody elektroluminescencyjne itd., opiera się na wykorzystaniu zjawisk zachodzących w stanie zakłóconej równowagi termodynamicznej.

W działaniu większości elementów półprzewodnikowych korzysta się z faktu, że w materiale półprzewodnikowym – w odróżnieniu np. od metalu – można za pomocą czynników zewnętrznych zmieniać w wybranych obszarach koncentracje nośników ładunku, czyli wytwarzać w nich lub wprowadzać do nich tzw. *nośniki nadmiarowe.* Po ustaniu działania tych czynników półprzewodnik powraca samorzutnie do stanu równowagi termodynamicznej. Proces ten, nazywany ogólnie *relaksacją*[1], może przebiegać w różny sposób, zależnie od rodzaju półprzewodnika i od warunków występujących w rozpatrywanym obszarze. Jednym z ważnych mechanizmów prowadzących do zaniku nośników nadmiarowych w półprzewodniku jest *rekombinacja* elektronów i dziur.

Zakłócenie równowagi termodynamicznej może polegać nie tylko na wytworzeniu w danym obszarze zwiększonych koncentracji nośników, lecz również na zmniejszeniu ich koncentracji w stosunku do odpowiednich wartości równo-

[1] *Relaksacją* nazywamy ogólnie proces ustalania się równowagi termodynamicznej w układzie wytrąconym z tego stanu. W książce tej będą rozpatrywane różne rodzaje procesów relaksacyjnych w półprzewodnikach. W niniejszym rozdziale słowo „relaksacja" będzie jednakże używane jedynie w odniesieniu do procesów rekombinacyjno-generacyjnych.

wagowych. W tym przypadku procesem, który po usunięciu czynnika zakłócającego przywraca stan równowagi termodynamicznej jest *generacja termiczna*. Oba te procesy zachodzą jednocześnie, dlatego też obejmiemy je wspólną nazwą *zjawisk rekombinacyjno-generacyjnych*. Jak zobaczymy dalej, w pewnych warunkach rekombinacja ma przewagę nad generacją termiczną, w innych natomiast zachodzi sytuacja odwrotna.

W rozdziale tym będą omówione najważniejsze mechanizmy zjawisk rekombinacyjno-generacyjnych w półprzewodnikach znajdujących zastosowanie w elementach półprzewodnikowych.

9.1 NOŚNIKI NADMIAROWE

Rozważmy jednorodny półprzewodnik, w którym koncentracje równowagowe elektronów i dziur wynoszą odpowiednio n_0 i p_0. Z definicji są to wielkości niezależne od czasu. Jeżeli zakłócimy równowagę termodynamiczną, wprowadzając lub wytwarzając nośniki nadmiarowe, to koncentracje te przyjmą wartości n i p. W rozdziale tym (z wyjątkiem p. 9.6) ograniczymy się do przypadku, w którym procesy rekombinacyjno-generacyjne przebiegają w sposób równomierny w całej objętości półprzewodnika, tzn. założymy, że koncentracje n i p są niezależne od współrzędnych rozpatrywanego punktu. Ogólnie rzecz biorąc, są one jednak funkcjami czasu. Dogodnie jest każdą z nich wyrażać w postaci sumy odpowiedniej koncentracji równowagowej i koncentracji nośników nadmiarowych. Można więc napisać

$$n(t) = n_0 + \Delta n(t) \tag{9/1}$$

$$p(t) = p_0 + \Delta p(t) \tag{9/2}$$

Wielkości Δn i Δp nazywamy krótko *koncentracjami nadmiarowymi*. Mogą one przyjmować zarówno wartości dodatnie, jak i ujemne. W pierwszym przypadku mamy do czynienia z rzeczywistym nadmiarem nośników w stosunku do warunków rownowagowych, w drugim zaś – z *niedoborem* nośników.

Zmiany koncentracji elektronów i dziur w stosunku do ich wartości równowagowych mogą następować w wyniku działania różnych czynników zewnętrznych i przebiegać w różny sposób. W elementach półprzewodnikowych istotne znaczenie mają następujące sposoby wytwarzania nośników nadmiarowych:

a) generacja par elektron-dziura wywołana absorpcją promieniowania elektromagnetycznego – zjawisko to nosi nazwę *wewnętrznego efektu fotoelektrycznego*;

b) generacja par elektron-dziura wywołana jonizacją atomów półprzewodnika w wyniku zderzeń nośników z siecią krystaliczną. Jest to proces odwrotny w stosunku do zjawiska Augera omawianego w p. 7.3. Ma on istotne znaczenie wówczas, gdy nośniki (elektrony lub dziury) poruszają się z dostatecznie dużymi prędkościami, a to zachodzi w warunkach, gdy w półprzewodniku występuje silne pole elektryczne. Zjawisko jonizacji zderzeniowej będzie rozpatrywane w p. 10.6.3);

c) *wstrzykiwanie* nośników mniejszościowych, czyli wprowadzanie tych nośników do danego obszaru półprzewodnika z przylegającego do niego innego obszaru (np. z metalu lub półprzewodnika). Może również zachodzić proces odwrotny, tj. wyciąganie nośników z danego obszaru do obszaru sąsiedniego. Proces ten nazywamy *ekstrakcją* nośników. Zagadnienia te będą szczegółowo analizowane w rozdz. 12.

W stanie zakłóconej równowagi termodynamicznej rozkład koncentracji nośników w półprzewodniku nie podlega statystyce Fermiego-Diraca. W tych warunkach nie można więc posługiwać się pojęciem poziomu Fermiego. Często jednak przyjmuje się, że w półprzewodniku, w którym występują nośniki nadmiarowe, poziom Fermiego ulega rozszczepieniu na 2 quasi-poziomy, z których jeden jest związany z koncentracją elektronów w pasmie przewodnictwa, drugi zaś z koncentracją dziur w pasmie walencyjnym. Zagadnienie to jest omówione w dodatku F.

9.2 REKOMBINACJA BEZPOŚREDNIA

Najprostszym rodzajem rekombinacji nośników jest tzw. *rekombinacja bezpośrednia*, nazywana również rekombinacją *międzypasmową*. Polega ona na *bezpośrednim* przejściu elektronu z pasma przewodnictwa na jeden z nieobsadzonych stanów w pasmie walencyjnym. W wyniku takiego przejścia znika 1 elektron w pasmie przewodnictwa i 1 dziura w pasmie walencyjnym, a więc ulega „anihilacji" 1 para elektron-dziura. Uwolniona przy tym energia zostaje uniesiona przez foton (rekombinacja promienista) albo też przekazana innemu elektronowi w pasmie przewodnictwa lub dziurze w pasmie walencyjnym (efekt Augera)[1].

Proces odwrotny polega na przejściu elektronu z pasma walencyjnego do pasma przewodnictwa. W wyniku tego zostaje kreowana (generowana) para elektron-dziura.

W obu tych procesach elektrony i dziury pojawiają się i znikają parami. Jeżeli zatem w rozpatrywanym obszarze półprzewodnika nie występuje pole elektryczne, to w każdej chwili jest spełniony warunek *neutralności elektrycznej*, wyrażony równością

$$\Delta n(t) = \Delta p(t) \qquad (9/3)$$

Ze wzoru (9/3) wynika, że zmiany koncentracji elektronów i dziur zachodzą z tą samą szybkością. Mogą więc być opisane równaniem

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \frac{\mathrm{d}p}{\mathrm{d}t} = G - R \qquad (9/4)$$

gdzie: G – *szybkość generacji*, wyrażająca liczbę par elektron-dziura genero-

[1] Energia wydzielająca się w czasie rekombinacji może być również uniesiona przez odpowiednio dużą ilość emitowanych fononów. Prawdopodobieństwo takich wielofononowych procesów jest jednakże bardzo małe, dlatego też procesy te zwykle nie odgrywają istotnej roli.

wanych w jednostce czasu w jednostce objętości półprzewodnika, R zaś – *szybkość rekombinacji* zdefiniowana w analogiczny sposób.

Generacja nośników jest wynikiem 2 niezależnych procesów zachodzących w półprzewodniku: generacji termicznej i generacji wywołanej działaniem czynników zewnętrznych, np. w postaci promieniowania elektromagnetycznego. Oznaczając szybkości tych procesów odpowiednio przez G_{th} i g, można szybkość generacji wyrazić w postaci sumy

$$G = G_{th}+g \tag{9/5}$$

Jak dalej zobaczymy, szybkość generacji termicznej G_{th} zależy od rodzaju półprzewodnika (między innymi od jego przerwy energetycznej W_g) i od temperatury, nie zależy natomiast od koncentracji nośników.

Stwierdzenie to może budzić pewne wątpliwości. Szybkość generacji termicznej powinna bowiem zależeć od liczby elektronów w pasmie walencyjnym (ta zaś zależy pośrednio od koncentracji dziur w tym pasmie) i od liczby nieobsadzonych stanów w pasmie przewodnictwa (tzn. pośrednio od koncentracji elektronów). Koncentracja elektronów w pasmie walencyjnym jest jednakże o kilka rzędów wielkości większa od największej osiągalnej koncentracji dziur i podobnie koncentracja nieobsadzonych stanów w pasmie przewodnictwa jest tak duża, że w zasadzie nie ulega ona zmianie wskutek obsadzenia części z nich przez elektrony. Tym tłumaczy się praktyczna niezależność szybkości generacji termicznej od koncentracji nośników.

Szybkość generacji nośników nadmiarowych g zależy natomiast od intensywności wywołującego ją bodźca zewnętrznego, np. od natężenia promieniowania elektromagnetycznego.

Z kolei szybkość rekombinacji R jest proporcjonalna do liczby elektronów mogących przejść z pasma przewodnictwa do pasma walencyjnego (tzn. do koncentracji elektronów n) i do liczby nieobsadzonych stanów w pasmie walencyjnym (tzn. do koncentracji dziur p).

Szybkość tę można więc wyrazić wzorem[1]

$$R = Cnp \tag{9/6}$$

gdzie C jest to tzw. *współczynnik rekombinacji*, zależny od rodzaju półprzewodnika i od temperatury. Wartości tego współczynnika dla kilku półprzewodników podano w tabl. 9/1.

Zazwyczaj nie interesuje nas całkowita szybkość rekombinacji lecz nadwyżka szybkości rekombinacji w warunkach nierównowagowych nad jej wartością w stanie równowagi termodynamicznej. Z prawa równowagi szczegółowej wynika, że ta ostatnia jest równa szybkości generacji termicznej G_{th}. Interesującą nas wielkością jest więc różnica $(R-G_{th})$. Wielkość tę będziemy oznaczali symbolem U_{rg} i nazywali *wypadkową szybkością procesów rekombinacyjno-generacyjnych*, lub krócej – *wypadkową szybkością rekombinacji.* Mamy zatem

$$U_{rg} = R-G_{th} \tag{9/7}$$

[1] Wzór ten dotyczy sytuacji, w której efekt Augera nie odgrywa istotnej roli. Rekombinacja Augera będzie omawiana w p. 9.5.

9/1 Wartości współczynnika rekombinacji C dla różnych półprzewodników w temperaturze 300 K

Półprzewodnik	Przerwa energetyczna	W_g	C
		eV	cm^3/s
GaAs	prosta	1,43	$7{,}21 \cdot 10^{-10}$
GaSb	prosta	0,70	$2{,}39 \cdot 10^{-10}$
InAs	prosta	0,35	$8{,}5 \cdot 10^{-11}$
InSb	prosta	0,18	$4{,}58 \cdot 10^{-11}$
GaP	skośna	2,27	$5{,}37 \cdot 10^{-14}$
Si	skośna	1,12	$1{,}79 \cdot 10^{-15}$
Ge	skośne	0,66	$5{,}25 \cdot 10^{-14}$

Źródło: Y. P. Varshni: Band-to-band radiative recombination in groups IV, VI and III-V semiconductors. *Phys. Stat. Sol.* **19**, 459 (1967) oraz **20**, 9 (1967).

Wielkość ta określa szybkość, z jaką półprzewodnik powraca do stanu równowagi termodynamicznej po ustaniu działania czynnika zakłócającego tę równowagę.

Korzystając ze wzorów (9/5) i (9/7) można zależność (9/4) przedstawić w postaci

$$\frac{dn}{dt} = \frac{dp}{dt} = g - U_{rg} \tag{9/8}$$

W warunkach równowagi termodynamicznej wszystkie te wielkości stają się równe zeru, a zatem $R = G_{th}$. Jednocześnie $n = n_0$, $p = p_0$, a dla półprzewodnika niezdegenerowanego $n_0 p_0 = n_i^2$. Biorąc to pod uwagę, otrzymuje się na podstawie poprzednio podanych wzorów następujące wyrażenie na szybkość generacji termicznej:

$$G_{th} = C n_i^2 \tag{9/9}$$

a więc, tak jak to już zaznaczono, szybkość ta zależy od W_g i od T.

Podstawiając (9/6) i (9/9) do wzoru (9/7), otrzymamy z kolei

$$U_{rg} = C\,(np - n_i^2) \tag{9/10}$$

Wypadkowa szybkość rekombinacji zależy więc od iloczynu np. Może ona przy tym przyjmować zarówno wartości dodatnie, jak i ujemne.

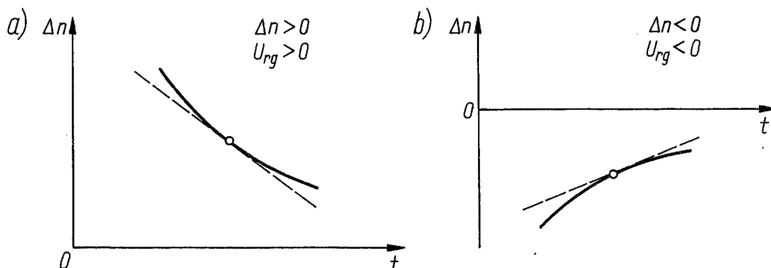
W przypadku nadmiaru nośników iloczyn np jest większy od n_i^2, a zatem U_{rg} jest dodatnie. Oznacza to, że procesy rekombinacyjne dominują nad generacją termiczną. Odwrotnie, w warunkach niedoboru nośników $np < n_i^2$, co sprawia, że U_{rg} staje się ujemne. Generacja termiczna ma więc teraz przewagę nad rekombinacją.

Ujemny znak wypadkowej szybkości rekombinacji nie powinien budzić zastrzeżeń, ponieważ poprzednio przyjęliśmy, że niedobór nośników będziemy traktowali jako występowanie ujemnej koncentracji nośników nadmiarowych.

Ze wzorów (9/8) i (9/1) wynika, że po ustaniu działania czynnika zakłócającego, powrót do stanu równowagi termodynamicznej zachodzi z szybkością

$$\frac{d(\Delta n)}{dt} = -U_{rg} \tag{9/11}$$

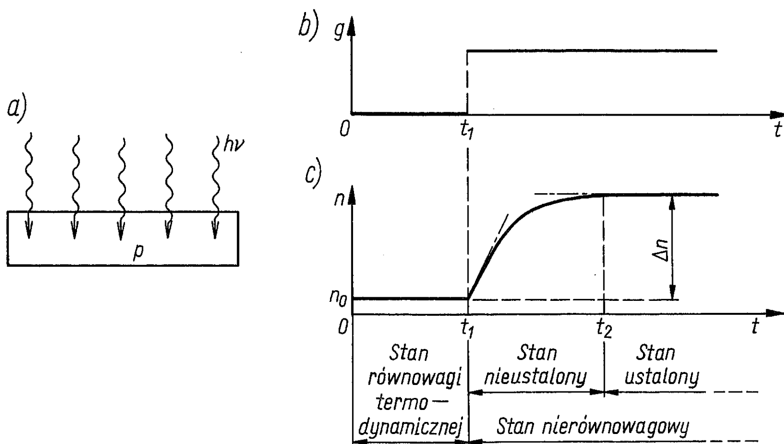
Jeżeli więc Δn jest dodatnie (rys. 9/1a), to U_{rg} jest także dodatnie, jeśli natomiast $\Delta n < 0$ (rys. 9/1b), to również $U_{rg} < 0$.



9/1 Proces zanikania nadmiarowych elektronów po ustaniu działania czynnika zakłócającego: a) $\Delta n > 0$ — rzeczywisty nadmiar elektronów; b) $\Delta n < 0$ — niedobór elektronów

Rozpatrzmy teraz półprzewodnik typu *p*, znajdujący się początkowo w stanie równowagi termodynamicznej, i załóżmy, że w chwili t_1 zostało włączone źródło promieniowania elektromagnetycznego wywołującego generację nośników nadmiarowych w tym półprzewodniku (rys. 9/2a). Zależność szybkości generacji *g* od czasu przedstawiono na rys. 9/2b. Poniżej podano przebieg zmian koncentracji nośników mniejszościowych, tzn. elektronów[1]. Początkowo wzrasta

[1] Zmiany koncentracji dziur są identyczne, ponieważ $\Delta p = \Delta n$. Jednakże, ze względu na to, że w półprzewodniku typu *p* zachodzi $p_0 \gg n_0$, względne zmiany koncentracji elektronów są o wiele rzędów wielkości większe od względnych zmian koncentracji dziur. Z tego względu interesujemy się tu bardziej nośnikami mniejszościowymi.



9/2 Generacja nośników nadmiarowych w półprzewodniku typu *p*: a) próbka półprzewodnika poddana działaniu promieniowania o energii fotonów $h\nu$; b) zależność szybkości generacji nośników nadmiarowych od czasu; c) zależność koncentracji nośników mniejszościowych (elektronów) od czasu

ona liniowo od wartości równowagowej n_0. Wraz z nią powiększa się szybkość rekombinacji U_{rg}, w wyniku czego maleje wyrażenie po prawej stronie równania (9/8), a zatem zmniejsza się szybkość zmian koncentracji elektronów. Po upływie dostatecznie długiego okresu czasu (praktycznie w chwili t_2) szybkość rekombinacji U_{rg} staje się równa szybkości generacji g. Oznacza to, że intensywność procesów przeciwdziałających zachodzącym zmianom osiągnęła taki poziom, przy którym ustala się nowy stan stacjonarny. Charakteryzuje się on tym, że parametry układu (w rozpatrywanym przypadku – koncentracje nośników) nie zmieniają się w czasie.

Należy zwrócić uwagę na istotną różnicę między omawianym tu stanem stacjonarnym a stanem równowagi termodynamicznej. W obu przypadkach parametry układu zachowują stałe wartości w czasie. Oba stany są więc stanami *ustalonymi*. Stan równowagi termodynamicznej charakteryzuje się jednak ponadto brakiem wszelkich przepływów, w tym również brakiem dopływu energii z zewnątrz, natomiast w stanie stacjonarnym mogą występować przepływy np. energii lub nośników, z tym tylko, że ich natężenia nie zmieniają się w czasie. W rozpatrywanym przypadku mamy do czynienia z dopływem energii promieniowania elektromagnetycznego, w wyniku czego powstają w półprzewodniku nośniki nadmiarowe. Układ jest więc wytrącony ze stanu równowagi termodynamicznej.

Zbadajmy zależność wypadkowej szybkości rekombinacji od stopnia zakłócenia równowagi termodynamicznej, tzn. od koncentracji Δn. W tym celu podstawmy wyrażenia (9/1) i (9/2) do wzoru (9/10). Otrzymamy wówczas

$$U_{rg} = C\,(n_0\,\Delta p + p_0\,\Delta n + \Delta n\,\Delta p) \tag{9/12}$$

a ponieważ $\Delta n = \Delta p$, zatem

$$U_{rg} = C\,[(n_0 + p_0)\,\Delta n + (\Delta n)^2] \tag{9/13}$$

Jest to zależność nieliniowa. Często jednak można założyć, że koncentracja nośników nadmiarowych jest mała w porównaniu z sumą koncentracji równowagowych, tzn. że jest spełniony warunek

$$\Delta n \ll n_0 + p_0 \tag{9/14}$$

Mówimy wówczas, że mamy do czynienia z tzw. *małym poziomem zakłócenia*. W tych warunkach można w przybliżeniu napisać

$$U_{rg} = C\,(n_0 + p_0)\,\Delta n \tag{9/15}$$

Wypadkowa szybkość rekombinacji jest więc w tym przypadku proporcjonalna do koncentracji nadmiarowej Δn. Zależność (9/15) można również przedstawić w postaci

$$U_{rg} = \frac{\Delta n}{\tau} \tag{9/16}$$

przy czym wprowadzony tu nowy parametr τ ma wymiar czasu i nosi nazwę *czasu życia* nośników nadmiarowych.

Podstawiając wzór (9/16) do (9/8), otrzymuje się następujące równanie różniczkowe opisujące zależność koncentracji nadmiarowej Δn od czasu przy małym poziomie zakłócenia

$$\frac{d(\Delta n)}{dt} = g - \frac{\Delta n}{\tau} \tag{9/17}$$

W stanie ustalonym pochodna $d(\Delta n)/dt$ znika. Wypływa stąd wniosek, że koncentracja nośników nadmiarowych

$$\Delta n = g\tau \tag{9/18}$$

Wzór ten tłumaczy nazwę „czas życia" nośników nadmiarowych nadaną parametrowi τ. Wynika z niego bowiem, że uśredniony na wszystkie nośniki nadmiarowe czas ich istnienia w półprzewodniku jest równy τ[1].

Problem 9.1. W próbce półprzewodnika występuje w stanie ustalonym koncentracja nadmiarowa Δn. W chwili $t = 0$ zostało wyłączone źródło promieniowania wywołującego generację tych nośników. Wyznacz przebieg zależności $\Delta n(t)$, zakładając, że mamy do czynienia z małym poziomem zakłócenia.

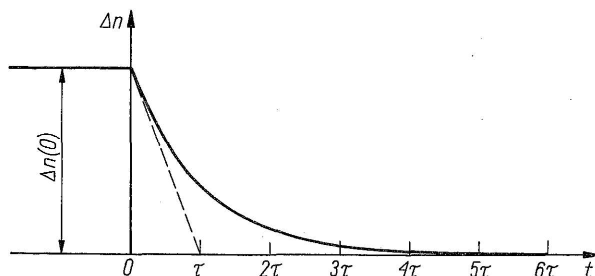
Rozwiązanie: Kładąc w równaniu (9/17) dla $t > 0$ wartość $g = 0$, otrzymujemy równanie

$$\frac{d(\Delta n)}{dt} = -\frac{\Delta n}{\tau} \tag{9/19}$$

którego rozwiązanie ma postać

$$\Delta n(t) = \Delta n(0) \exp\left(-\frac{t}{\tau}\right) \tag{9/20}$$

Widzimy więc, że jeżeli odchylenie od stanu równowagi termodynamicznej jest dostatecznie małe, to proces relaksacji (czyli powrót do stanu równowagi) przebiega wg wykładniczej funkcji czasu (rys. 9/3), przy czym stała czasowa tego procesu jest równa czasowi życia nośników nadmiarowych τ. Po czasie τ koncentracja nośników nadmiarowych maleje e-krotnie, a po czasie 5τ spada do wartości mniejszej niż 1% wartości początkowej. Zwykle można więc uważać, że po upływie tego czasu zakłócenie stanu równowagi termodynamicznej przestaje istnieć.



9/3 Zanikanie koncentracji nośników nadmiarowych po ustaniu działania czynnika zakłócającego — mały poziom zakłócenia

Problem 9.2. Z rys. 9/3 wynika, że nośniki nadmiarowe zanikają w półprzewodniku w sposób stopniowy, a więc czas ich istnienia jest dla poszczególnych nośników różny. Wykaż, że czas ten uśredniony na wszystkie nośniki nadmiarowe jest równy czasowi życia τ.

---

[1] Wzór (9/18) dotyczy dowolnej populacji w stanie ustalonym. Jeżeli w danym obszarze pojawia się (np. rodzi się) w jednostce czasu g osobników, a ich średni czas życia wynosi τ, to liczba osobników w tym obszarze jest równa $g\tau$.

Ze wzorów (9/15) i (9/16) wynika, że przy małym poziomie zakłócenia czas życia nośników nadmiarowych wyraża się następującym ogólnym wzorem:

$$\tau = \frac{1}{C(n_0+p_0)} \tag{9/21}$$

słusznym zarówno dla półprzewodnika samoistnego, jak i dla domieszkowego (typu n lub p). Wynika z niego, że czas ten jest tym krótszy, im silniej domieszkowany jest dany półprzewodnik.

Dla półprzewodnika typu p wzór (9/21) można sprowadzić do postaci

$$\tau = \frac{1}{CN_a} \tag{9/22}$$

Jeżeli poziom zakłócenia stanu równowagi termodynamicznej jest duży, tzn. nierówność (9/14) nie jest spełniona, to równanie opisujące proces relaksacji ma postać

$$\frac{d(\Delta n)}{dt} = -C[(n_0+p_0)\Delta n+(\Delta n)^2] \tag{9/23}$$

Równanie to może być dość łatwo rozwiązane, uzyskany wynik jest jednak mało przejrzysty. Dlatego też rozpatrzymy tu drugi skrajny przypadek, tj. bardzo dużego poziomu zakłócenia, który występuje wówczas, gdy

$$\Delta n \gg n_0+p_0 \tag{9/24}$$

Równanie (9/23) upraszcza się wówczas do postaci

$$\frac{d(\Delta n)}{dt} = -C(\Delta n)^2 \tag{9/25}$$

a jego rozwiązaniem jest funkcja

$$\Delta n(t) = \frac{\Delta n(0)}{1+\Delta n(0)\,Ct} \tag{9/26}$$

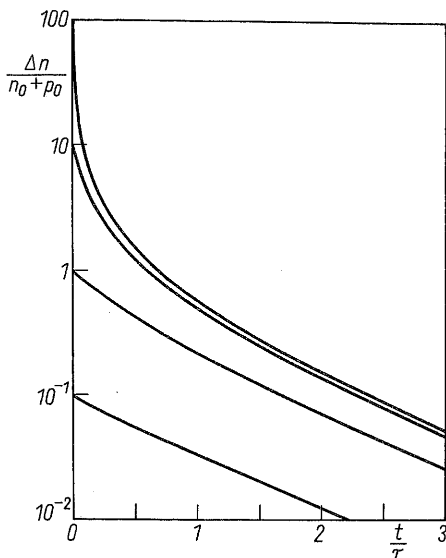
Zanik nośników nadmiarowych zachodzi teraz, jak widać, wg hiperbolicznej funkcji czasu.

Przebieg zależności $\Delta n(t)$ po wyłączeniu czynnika zakłócającego przedstawiono dla różnych poziomów zakłócenia na rys. 9/4. Na osi rzędnych zastosowano tu podziałkę logarytmiczną. W tych warunkach wykresem funkcji wykładniczej jest linia prosta. Jak widzimy, przy małym poziomie zakłócenia relaksacja przebiega wg funkcji wykładniczej zgodnie ze wzorem (9/20).

W miarę zwiększania poziomu zakłócenia obserwuje się coraz większe odchylenie funkcji $\Delta n(t)$ od przebiegu wykładniczego. W tych warunkach pojęcie czasu życia nośników staje się niejednoznaczne, dlatego też będziemy się nim posługiwali tylko wtedy, kiedy zanik nośników nadmiarowych przebiega wg wykładniczej funkcji czasu. Tylko wówczas można bowiem czas życia nośników traktować jako stałą czasową przebiegu relaksacyjnego[1].

---

[1] Niektórzy autorzy posługują się pojęciem czasu życia nośników także przy dużym poziomie zakłócenia, definiując go jako współczynnik τ we wzorze (9/16). Tak określony „czas życia" zależy od chwilowej wartości koncentracji nadmiarowej Δn, a więc zmienia się w miarę tego, jak koncentracja ta maleje w wyniku rekombinacji nośników.



9/4
Zanikanie koncentracji nośników nadmiarowych przy różnych poziomach zakłócenia
τ — czas życia nośników nadmiarowych przy małym poziomie zakłócenia

Na rysunku 9/4 można zaobserwować, że im wyższy jest poziom zakłócenia, tym szybciej zachodzi proces relaksacji.

Dla półprzewodników z prostą przerwą energetyczną, w których prawdopodobieństwo przejść międzypasmowych jest duże, współczynnik rekombinacji C przyjmuje stosunkowo duże wartości (zob. tabl. 9/1), a wskutek tego czas życia nośników nadmiarowych jest bardzo krótki. Przykładowo, dla GaAs przy małych koncentracjach domieszek czas ten jest rzędu 10^{-8}s.

Z kolei dla półprzewodników ze skośną przerwą energetyczną np. dla germanu i krzemu, bezpośrednie przejścia międzypasmowe są tak mało prawdopodobne, że rekombinacja bezpośrednia nie odgrywa istotnej roli. Procesy rekombinacyjno-generacyjne zachodzą tu w inny sposób, o czym będzie mowa w następnym punkcie.

Pomimo ograniczonego znaczenia rekombinacji bezpośredniej, omówieniu tego zjawiska poświęciliśmy dość dużo miejsca, a to z tego względu, że wprowadzone tu pojęcia i uzyskane zależności znajdą zastosowanie również przy rozpatrywaniu innych rodzajów rekombinacji.

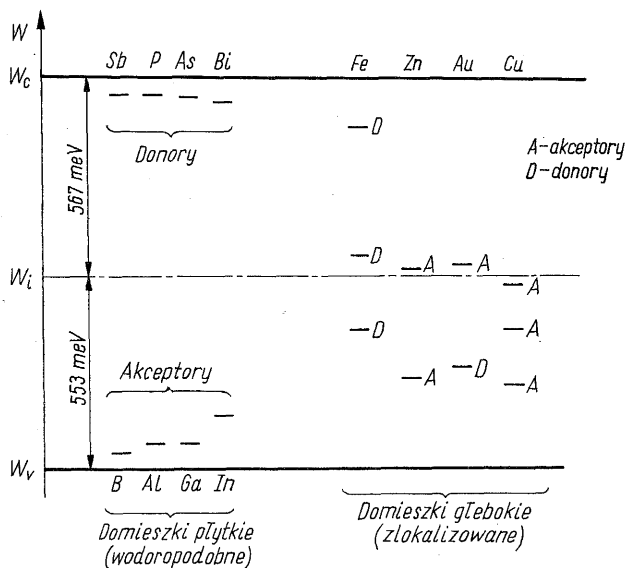
9.3 REKOMBINACJA POŚREDNIA

9.3.1 Pułapki w półprzewodniku

Rekombinacja *pośrednia* jest to proces, w którym przejście elektronu z pasma przewodnictwa do pasma walencyjnego odbywa się w sposób dwustopniowy poprzez pewne zlokalizowane stany kwantowe, nazywane *pułapkami.* Są one utworzone przez różne rodzaje defektów struktury krystalicznej, o których była mowa w p. 4.5.

Z pułapkami są związane dyskretne poziomy energetyczne w przerwie energetycznej półprzewodnika. Na rysunku 9/5 pokazano na przykładzie krzemu

poziomy wytwarzane przez różne rodzaje domieszek. W lewej części rysunku zestawiono tzw. *domieszki płytkie*, które mają zasadniczy wpływ na koncentracje równowagowe nośników, ale których udział w procesach rekombinacyjnych jest nieistotny. Prawa część rysunku obejmuje natomiast tzw. *domieszki głębokie*, które ze względu na dużą wartość energii jonizacji praktycznie nie mają wpływu na koncentracje równowagowe nośników. Pełnią one natomiast rolę pułapek ułatwiających dwustopniowe przechodzenie elektronów z pasma przewodnictwa do pasma walencyjnego lub w kierunku odwrotnym.



9/5 Poziomy domieszkowe wytwarzane przez różne pierwiastki w krzemie

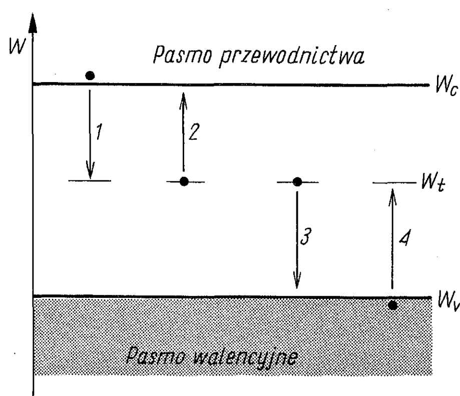
Jak wynika z rys. 9/5 niektóre domieszki wytwarzają kilka stanów pułapkowych, przy czym jedne z nich mogą mieć charakter *donorowy* (tzn. w stanie obsadzonym przez elektron są elektrycznie obojętne), inne zaś – charakter *akceptorowy* (w stanie obsadzonym przez elektron są zjonizowane). Podobne stany pułapkowe są wytwarzane przez inne defekty strukturalne, np. luki węzłowe.

Na rysunku 9/6 przedstawiono model pasmowy półprzewodnika, w którym występują pułapki, przy czym z każdą z nich jest związany pojedynczy poziom W_t, leżący wewnątrz przerwy energetycznej[1]. Rozpatrując procesy rekombinacyjno-generacyjne należy uwzględnić 4 rodzaje przejść elektronowych, które mogą występować w tym półprzewodniku[2]. Pokazano je na rys. 9/6.

Elektron z pasma przewodnictwa może przejść na nieobsadzony poziom pułapkowy (przejście *1*) – proces ten nazywamy *wychwytem elektronu* z pasma

[1] Indeks *t* w symbolu W_t pochodzi od angielskiego słowa „trap" (pułapka).

[2] Nie bierzemy tu pod uwagę przejść międzypasmowych, które były rozpatrywane w poprzednim punkcie.



9/6
Przejścia elektronowe przy rekombinacji pośredniej: *1* — wychwyt elektronu z pasma przewodnictwa; *2* — emisja elektronu do pasma przewodnictwa; *3* — wychwyt dziury z pasma walencyjnego; *4* — emisja dziury do pasma walencyjnego

przewodnictwa. W wyniku tego procesu elektron zostaje zlokalizowany w stanie pułapkowym, a jego energia spada z wartości, jaką miał w pasmie przewodnictwa, do wartości W_t. Wyzwolona przy tym nadwyżka energii zostaje najczęściej zużyta na emisję pewnej liczby fononów.

Elektron schwytany w pułapkę może zostać z powrotem wzbudzony do pasma przewodnictwa (przejście *2*) albo przejść do pasma walencyjnego (przejście *3*). Pierwszy proces nazywamy *emisją elektronu* z poziomu pułapkowego do pasma przewodnictwa, drugi zaś można traktować jako *wychwyt dziury* z pasma walencyjnego. Procesem odwrotnym jest *emisja dziury* z poziomu pułapkowego do pasma walencyjnego, co jest równoważne przejściu elektronu z tego pasma na poziom pułapkowy (przejście *4*).

Każdy z tych 4 procesów zachodzi z pewną szybkością, określającą liczbę przejść przypadających na jednostkę objętości półprzewodnika i jednostkę czasu. Szybkości te zależą od koncentracji pułapek, ich własności fizycznych (m.in. od położenia poziomu W_t w stosunku do krawędzi pasm energetycznych) i od koncentracji nośników w poszczególnych pasmach.

W stanie równowagi termodynamicznej każdy proces jest kompensowany przez swój proces odwrotny, a zatem przejścia *1* i *2* oraz *3* i *4* zachodzą (statystycznie biorąc) z odpowiednio równymi szybkościami. W wyniku tego szybkość rekombinacji R jest równa szybkości generacji termicznej G_{th}.

W warunkach, w których występuje nadmiar elektronów i dziur, przejścia *1* i *3* zachodzą z większą intensywnością niż przejścia *2* i *4*. Szybkość rekombinacji przewyższa teraz szybkość generacji termicznej, co sprawia, że rekombinacja staje się procesem przywracającym stan równowagi termodynamicznej w półprzewodniku.

Odwrotnie, w warunkach niedoboru nośników, przejścia *1* i *3* stają się mniej prawdopodobne od przejść *2* i *4*. Procesem dominującym staje się więc teraz generacja termiczna.

9.3.2 Rekombinacja SRH

Najprostsze równania opisujące zjawisko rekombinacji pośredniej zostały sformułowane przez Shockley'a i Reada oraz Halla. Przyjęty przez nich model opiera się na założeniu, że w półprzewodniku występują stany pułapkowe jednego rodzaju o koncentracji N_t i że z każdym z nich jest związany pojedynczy poziom energetyczny W_t. Rekombinacja zachodząca wg tego schematu nosi nazwę *rekombinacji Shockley'a-Reada-Halla* lub w skrócie rekombinacji SRH.

W celu wyznaczenia wypadkowej szybkości rekombinacji, należy wyznaczyć szybkości, z jakimi zachodzą poszczególne procesy pokazane na rys. 9/6. Oznaczając je odpowiednio symbolami $r_1 \dots r_4$, możemy napisać następujące podwójne równanie:

$$U_{rg} = r_1 - r_2 = r_3 - r_4 \tag{9/27}$$

Wynika ono z faktu, że zarówno różnica $r_1 - r_2$, jak i $r_3 - r_4$ wyznaczają liczbę elektronów przeniesionych w jednostce objętości i jednostce czasu z pasma przewodnictwa do pasma walencyjnego.

Szybkość wychwytu elektronów z pasma przewodnictwa jest proporcjonalna do koncentracji elektronów n w tym pasmie i do koncentracji pułapek nie obsadzonych przez elektrony. Oznaczając prawdopodobieństwo obsadzenia pułapki symbolem $f(W_t)$, można szybkość r_1 wyrazić wzorem

$$r_1 = C_n n N_t [1 - f(W_t)] \tag{9/28}$$

gdzie C_n jest współczynnikiem proporcjonalnym do prawdopodobieństwa odpowiedniego przejścia elektronowego.

W warunkach równowagi termodynamicznej funkcja $f(W_t)$ jest identyczna z funkcją rozkładu Fermiego-Diraca, którą dla odróżnienia będziemy tu oznaczali symbolem $f_F(W_t)$. W warunkach nierównowagowych natomiast obsadzenie stanów pułapkowych odbiega od wartości wynikającej z rozkładu Fermiego-Diraca, a zatem $f(W_t) \neq f_F(W_t)$. Jak dalej zobaczymy, znajomość funkcji $f(W_t)$ nie jest jednakże potrzebna do wyznaczenia poszukiwanej wielkości U_{rg}.

Szybkość emisji elektronów do pasma przewodnictwa jest z kolei proporcjonalna do koncentracji obsadzonych stanów pułapkowych i do efektywnej gęstości stanów w pasmie przewodnictwa N_c. Wynosi ona zatem

$$r_2 = C_n' N_c N_t f(W_t) \tag{9/29}$$

gdzie C_n' jest współczynnikiem analogicznym jak C_n, ale dla przejść zachodzących w przeciwnym kierunku.

Współczynniki C_n i C_n' są wzajemnie powiązane. Ze względu na to, że w stanie równowagi termodynamicznej $r_1 = r_2$, z przyrównania wyrażeń (9/28) i (9/29) wynika bowiem następująca równość:

$$C_n' N_c f_F(W_t) = C_n n_0 [1 - f_F(W_t)] \tag{9/30}$$

W równaniu tym zamiast n podstawiono n_0, a zamiast $f(W_t)$ napisano $f_F(W_t)$, ponieważ równanie to dotyczy stanu równowagi termodynamicznej.

Podstawiając

$$f_F(W_t) = \frac{1}{1+\exp\frac{W_t - W_F}{kT}} \qquad (9/31)$$

otrzymamy szukany związek między współczynnikami C_n i C_n'

$$C_n' = C_n \exp\left(-\frac{W_c - W_t}{kT}\right) \qquad (9/32)$$

Stosunek prawdopodobieństw przejść oznaczonych na rys. 9/6 symbolami *1* i *2* zależy więc od temperatury półprzewodnika i od tego, jak daleko od dna pasma przewodnictwa znajduje się poziom pułapkowy W_t. Prawdopodobieństwo emisji elektronu schwytanego w stanie pułapkowym maleje mianowicie w miarę tego, jak poziom W_t oddala się od W_c.

Analogiczne wzory określające szybkość wychwytu dziur z pasma walencyjnego i szybkość ich emisji do tego pasma mają postać

$$r_3 = C_p\, p\, N_t f(W_t) \qquad (9/33)$$

$$r_4 = C_p'\, N_v\, N_t\, [1 - f(W_t)] \qquad (9/34)$$

a z ich przyrównania w stanie równowagi termodynamicznej otrzymuje się

$$C_p' = C_p \exp\left(-\frac{W_t - W_v}{kT}\right) \qquad (9/35)$$

Podstawiając otrzymane wyrażenia do (9/27), można wyeliminować nieznaną funkcję $f(W_t)$ i po pewnych przekształceniach otrzymuje się następujący wzór na wypadkową szybkość rekombinacji:

$$U_{rg} = \frac{C_n\, C_p\, N_t(np - n_i^2)}{C_n\left(n + n_i \exp\frac{W_t - W_i}{kT}\right) + C_p\left(p + n_i \exp\frac{W_i - W_t}{kT}\right)} \qquad (9/36)$$

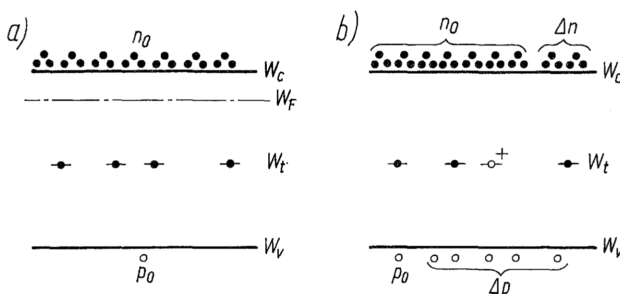
Wzór ten, jak widzimy, jest znacznie bardziej skomplikowany od poznanego wcześniej wzoru (9/10), dotyczącego rekombinacji bezpośredniej. Tam bowiem współczynnik rekombinacji C był dla danego półprzewodnika w określonej temperaturze wielkością stałą. Tu, natomiast, wyrażenie $(np - n_i^2)$ jest pomnożone przez wielkość zależną nie tylko od rodzaju półprzewodnika i temperatury, lecz również od koncentracji elektronów i dziur. Ze względu na skomplikowany charakter zależności (9/36), ograniczymy nasze rozważania do łatwiejszych, a jednocześnie ważniejszych przypadków.

9.3.3* Warunek neutralności elektrycznej

Przy analizowaniu rekombinacji bezpośredniej mieliśmy do czynienia z przypadkiem, w którym pojawieniu się elektronu w pasmie przewodnictwa (lub zniknięciu elektronu z tego pasma) towarzyszyło jednoczesne pojawienie się (zniknięcie)

dziury w pasmie walencyjnym. Konsekwencją tej jednoczesności był wzór (9/3), wyrażający warunek neutralności elektrycznej.

Nasuwa się obecnie pytanie, czy wzór ten może być stosowany również w przypadku rekombinacji pośredniej. Ścisła odpowiedź na to pytanie jest negatywna, ponieważ ze zmianą koncentracji nośników zmienia się wartość funkcji $f(W_t)$, wyrażającej prawdopodobieństwo obsadzenia stanów pułapkowych. Zmiany te przedstawiono w sposób poglądowy na rys. 9/7. Pokazano tam model pasmowy półprzewodnika typu *n* w warunkach równowagi termodynamicznej i w warunkach, w których występują nośniki nadmiarowe. Założono przy tym, że w pobliżu środka przerwy energetycznej półprzewodnika znajdują się zlokalizowane poziomy pułapkowe o energii W_t. Przyjmijmy, że są to pułapki typu donorowego. W stanie obsadzonym przez elektrony są one więc elektrycznie obojętne.



9/7 Stan obsadzenia poziomów pułapkowych w warunkach równowagowych (a) i nierównowagowych (b)

W stanie równowagi termodynamicznej (rys. 9/7a), praktycznie rzecz biorąc, wszystkie stany pułapkowe są obsadzone przez elektrony. Wynika to z faktu, że w rozpatrywanym półprzewodniku poziom Fermiego jest przesunięty w kierunku pasma przewodnictwa, a więc znajduje się znacznie wyżej niż poziom W_t. W półprzewodniku występują 4 rodzaje ładunków elektrycznych: nośniki oraz zjonizowane domieszki (nie uwidocznione na rysunku), a warunek neutralności elektrycznej wyraża się wzorem (8/37).

Po zakłóceniu równowagi termodynamicznej, np. wskutek oświetlenia półprzewodnika, pojawiają się nadmiarowe elektrony o koncentracji Δn i nadmiarowe dziury o koncentracji Δp. Jeżeli Δn jest małe w stosunku do n_0, to koncentracja elektronów zostaje zakłócona w niewielkim stopniu, a zatem, jak wynika ze wzoru (9/28), szybkość wychwytu elektronów r_1 prawie nie ulega zmianie. Szybkość wychwytu dziur – wzór (9/33) – powiększa się natomiast znacznie, ponieważ ich koncentracja może wzrosnąć np. wiele milionów razy. Wskutek powiększenia szybkości wychwytu dziur zmniejsza się obsadzenie stanów pułapkowych – zob. rys. 9/7b. Ponieważ, jak założyliśmy, są to pułapki typu donorowego, zatem w stanie nie obsadzonym przez elektrony są one dodatnio zjonizo-

wane. Oznaczając ich koncentrację symbolem Δm, możemy warunek neutralności elektrycznej półprzewodnika sformułować następująco:

$$\Delta n(t) = \Delta p(t) + \Delta m(t) \tag{9/37}$$

W rozpatrywanym półprzewodniku w stanie równowagi termodynamicznej $\Delta m = 0$. Po włączeniu źródła promieniowania wywołującego generację nośników nadmiarowych wszystkie 3 koncentracje występujące we wzorze (9/37) wzrastają monotonicznie, przy czym w każdej chwili jest spełniona równość wyrażona tym wzorem. Coraz większa ilość ładunków dodatnich zostaje więc związana ze stanami pułapkowymi. Zjawisko to nazywamy *pułapkowaniem* nośników – w rozpatrywanym przypadku mamy do czynienia z pułapkowaniem dziur, ponieważ stan pułapkowy nie obsadzony przez elektron można traktować jako obsadzony przez dziurę.

Po usunięciu źródła promieniowania koncentracje nośników oraz Δm powracają stopniowo do swych wartości równowagowych.

Problem 9.3. Dla rozpatrywanego wyżej półprzewodnika wyprowadź wzór na stopień obsadzenia stanów pułapkowych przez dziury. Załóż, że $W_t = W_i$ oraz $C_p = C_n$ i przyjmij, że we wzorze (9/37) Δm jest pomijalnie małe wobec Δp. Przedyskutuj otrzymany wzór w zależności od Δp, uwzględniając przypadki graniczne $\Delta p \to 0$ oraz $\Delta p \to \infty$.
Wskazówka: W celu wyznaczenia funkcji $f(W_t)$, należy wykorzystać równość $r_1 - r_2 = r_2 - r_4$, stanowiącą część zależności (9/27).

Problem 9.4. Przeprowadź dyskusję zjawiska pułapkowania nośników dla półprzewodnika typu n, zawierającego pułapki typu akceptorowego, i dla półprzewodnika typu p, zawierającego pułapki typu donorowego. W obu przypadkach załóż, że $W_t \approx W_i$.

9.3.4 Mały poziom zakłócenia

Przystąpimy teraz do dyskusji wzoru (9/36), przy czym założymy, że:

1) rozpatrywany półprzewodnik, np. typu n, jest dostatecznie silnie domieszkowany, tzn. $n_0 \gg n_i$;

2) poziom zakłócenia stanu równowagi jest mały, tzn. $\Delta p \ll n_0$;

3) omawiane w p. 9.3.3 zjawisko pułapkowania nośników nie ma istotnego wpływu na koncentracje nośników nadmiarowych, tzn. $\Delta m \ll \Delta p$.

Opierając się na tych założeniach, można wzór (9/36) uprościć do postaci

$$U_{rg} = \frac{C_n C_p N_t n_0 \Delta p}{C_n \left(n_0 + n_i \exp \dfrac{W_t - W_i}{kT}\right) + C_p n_i \exp \dfrac{W_i - W_t}{kT}} \tag{9/38}$$

Ze wzoru tego wynika, że wypadkowa szybkość rekombinacji osiąga duże wartości tylko wówczas, gdy poziom pułapkowy W_t znajduje się blisko poziomu samoistnego W_i, który, jak pamiętamy, leży prawie dokładnie w połowie przerwy energetycznej. Takie stany pułapkowe odznaczające się dużą skutecznością w procesach rekombinacyjno-generacyjnych nazywamy *centrami rekombinacyjnymi*. Z rysunku 9/5 wynika, że centra takie mogą być wytworzone w krzemie przez domieszkowanie tego półprzewodnika miedzią, złotem, żelazem i innymi metalami.

W miarę oddalania się poziomu W_t od W_i w jednym, bądź drugim kierunku, wzrasta szybko jeden z czynników wykładniczych w mianowniku wyrażenia (9/38) i wskutek tego zmniejsza się znacznie szybkość rekombinacji. Przy dużych wartościach $|W_t - W_i|$, a więc np. dla poziomów położonych w pobliżu dna pasma przewodnictwa lub wierzchołka pasma walencyjnego jest ona tak mała, że jak to już zaznaczono, obecność tych poziomów nie ma istotnego wpływu na procesy rekombinacyjno-generacyjne.

Zakładając, że różnica $|W_t - W_i|$ nie przekracza kilkudziesięciu meV, a C_p różni się od C_n nie więcej niż o 1 rząd wielkości, można wzór (9/38) uprościć dalej i otrzymać

$$U_{rg} = C_p N_t \Delta p \tag{9/39}$$

Wypadkowa szybkość rekombinacji jest więc w tych warunkach proporcjonalna do prawdopodobieństwa wychwytu nośników *mniejszościowych*, tzn. w rozpatrywanym tu półprzewodniku typu n – do współczynnika C_p, od którego zależy prawdopodobieństwo wychwytu dziury z pasma walencyjnego.

Wniosek ten jest intuicyjnie zrozumiały, ponieważ w półprzewodniku typu n, przy małym poziomie zakłócenia, prawie wszystkie stany pułapkowe są obsadzone przez elektrony (zob. rys. 9/7). W tych warunkach szybkość rekombinacji zależy od szybkości, z jaką elektrony są przenoszone z tych stanów do pasma walencyjnego, czyli od szybkości wychwytu dziur. Ponadto, co jest oczywiste, szybkość U_{rg} jest proporcjonalna do koncentracji pułapek N_t.

Ze wzorów (9/38) i (9/39) wynika dalej, że przy małym poziomie zakłócenia powrót do stanu równowagi termodynamicznej odbywa się z szybkością proporcjonalną do odchylenia od tego stanu, tzn. proporcjonalną do koncentracji nośników nadmiarowych. Relaksacja przebiega zatem wg wykładniczej funkcji czasu ze stałą czasową, która zgodnie ze wzorem (9/39) wynosi

$$\boxed{\tau_p = \frac{1}{C_p N_t}} \tag{9/40}$$

Wielkość tę nazywamy zwykle czasem życia nośników *mniejszościowych*, dla podkreślenia faktu, że jest ona zależna od prawdopodobieństwa wychwytu tych właśnie nośników.

Analogiczne rozumowanie dla półprzewodnika typu p prowadzi do wzorów

$$U_{rg} = C_n N_t \Delta n \tag{9/41}$$

$$\boxed{\tau_n = \frac{1}{C_n N_t}} \tag{9/42}$$

9.3.5* Duży poziom zakłócenia

Jeżeli koncentracja nośników nadmiarowych nie jest pomijalnie mała w stosunku do koncentracji nośników większościowych, jeśli np. dla półprzewodnika typu n warunek $\Delta p \ll n_0$ nie jest spełniony, to wypadkowa szybkość rekombinacji nie jest proporcjonalna do Δp i w wyniku tego, proces relaksacji nie zachodzi wg wykładniczej funkcji czasu. W tych warunkach, podobnie jak przy rekombinacji bezpośredniej, pojęcie czasu życia nośników staje się niejednoznaczne.

Jedynie przy bardzo dużym poziomie zakłócenia, gdy spełniony jest warunek $\Delta p \gg n_0 + p_0$, wzór (9/36) upraszcza się do postaci

$$U_{rg} = \frac{C_n C_p N_t}{C_n + C_p} \Delta p \qquad (9/43)$$

a zatem relaksacja przebiega znów wg funkcji wykładniczej ze stałą czasową, którą często oznacza się symbolem τ_∞. Jak wynika ze wzoru (9/43), wynosi ona

$$\tau_\infty = \left(\frac{1}{C_p} + \frac{1}{C_p}\right) \frac{1}{N_t} \qquad (9/44)$$

przy czym zależność ta dotyczy zarówno półprzewodnika typu n, jak i typu p. Przy znacznie różniących się wartościach współczynników C_n i C_p czas ten może w istotny sposób odbiegać od czasu życia nośników nadmiarowych przy małym poziomie zakłócenia.

9.4 CZAS ŻYCIA NOŚNIKÓW NADMIAROWYCH

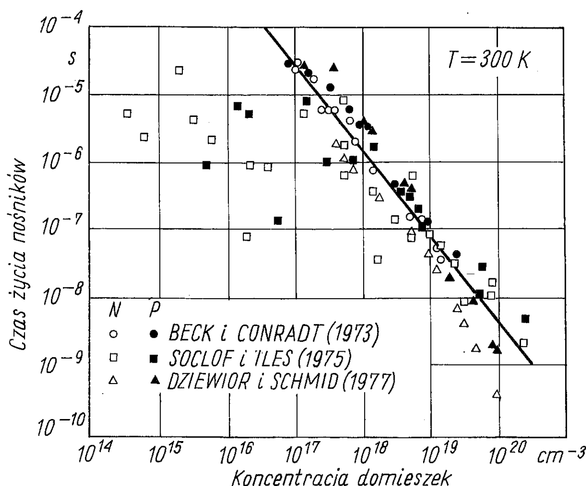
W poprzednich punktach stwierdziliśmy, że przy małym poziomie zakłócenia równowagi termodynamicznej proces relaksacji zachodzący w półprzewodniku po usunięciu czynnika zakłócającego tę równowagę przebiega wg wykładniczej funkcji czasu. Zachodzi on zatem z pewną stałą czasową, którą nazywamy *czasem życia nośników nadmiarowych.* Czas ten jest jedną z podstawowych wielkości charakteryzujących dany materiał i określających jego przydatność w różnych rodzajach elementów półprzewodnikowych.

Czas życia nośników może zmieniać się w szerokich granicach w zależności od materiału podstawowego (istotne różnice występują tu zależnie od tego, czy rozpatrywany półprzewodnik ma prostą, czy też skośną przerwę energetyczną) oraz od rodzaju i koncentracji występujących w nim centrów rekombinacyjnych. To z kolei zależy od przeprowadzonych procesów technologicznych i innych oddziaływań zewnętrznych (np. mechanicznych), w wyniku których powstają w krysztale defekty strukturalne.

Spośród różnych możliwych procesów technologicznych, którym poddawany jest kryształ w czasie wytwarzania elementu półprzewodnikowego, na szczególną uwagę zasługuje dyfuzja domieszek. Nawet w przypadku, gdy są to płytkie domieszki, które jak mówiliśmy, nie odgrywają istotnej roli w procesach rekombinacyjno-generacyjnych, to jednak, wnikając w głąb półprzewodnika,

wywołują one na swej drodze uszkodzenia sieci krystalicznej w postaci luk i przemieszczonych atomów. Te zakłócenia regularności sieci pełnią w półprzewodniku rolę centrów rekombinacyjnych determinujących czas życia nośników.

Uzyskiwane praktycznie wartości czasu życia nośników w krzemie w zależności od koncentracji domieszek zestawiono na rys. 9/8. Przy koncentracjach mniejszych niż 10^{17} cm^{-3} wartości te zawierają się zazwyczaj w granicach od 0,5 do 20 µs. Przy większych koncentracjach domieszek maleją one ze wzrostem N, w przybliżeniu proporcjonalnie do $N^{-1,34}$.



9/8 Zależność czasu życia nośników nadmiarowych w krzemie od koncentracji domieszek: N — dane doświadczalne dotyczące krzemu typu *n*, P — dane doświadczalne dotyczące krzemu typu *p*
(Zaczerpnięto z: R. M. Warner, Jr., B. L. Grung: *Transistors, fundamentals for the integrated-circuit engineer*. J. Wiley, New York 1983)

Pożądane wartości czasu życia zależą od rodzaju i przeznaczenia elementu półprzewodnikowego. Przykładowo, w fotodetektorach i niektórych rodzajach pamięci półprzewodnikowych są wymagane duże wartości czasu życia nośników. Przy produkcji tych elementów dąży się zatem do uzyskania możliwie „idealnej" struktury krystalicznej. W przypadku krzemu udaje się wówczas uzyskać czasy życia nośników dochodzące do kilku milisekund.

W diodach i tranzystorach typowe wartości czasu życia nośników powinny, w zależności od przeznaczenia, zawierać się w granicach do ok. 0,1 µs do kilkudziesięciu mikrosekund. Wartości takie uzyskuje się zwykle samorzutnie w wyniku procesów dyfuzji donorów i akceptorów, zapewniających potrzebne koncentracje nośników w poszczególnych obszarach danego elementu.

W niektórych przypadkach, np. w diodach przełączających o dużych szybkościach działania, pożądane są jak najmniejsze wartości czasu życia nośników. Jednym ze sposobów zapewnienia krótkich czasów życia jest wprowadzenie domieszek wytwarzających poziomy pułapkowe w pobliżu połowy przerwy energetycznej. Typowym przykładem takiej domieszki dla krzemu i germanu

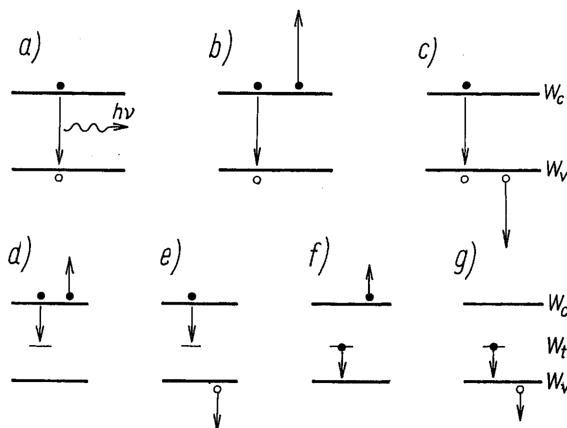
jest złoto. Dla krzemu typu n domieszkowanego złotem współczynnik C_p występujący we wzorze (9/40) wynosi ok. $5 \cdot 10^{-8}$ cm³/s. Oznacza to, że wprowadzając atomy złota o koncentracji N_t zmieniającej się w zakresie $2 \cdot 10^{14} - 2 \cdot 10^{17}$ cm^{-3}, można uzyskiwać czasy życia dziur w granicach od 10^{-7} do 10^{-10} s.

Innym sposobem zmniejszania czasu życia nośników jest napromieniowanie półprzewodnika cząstkami o dużej energii. Wywołuje to także uszkodzenia sieci krystalicznej, wskutek czego pojawiają się poziomy pułapkowe w przerwie energetycznej. Przykładowo, bombardowanie krzemu za pomocą elektronów wywołuje pojawienie się poziomu akceptorowego w odległości 0,4 eV od wierzchołka pasma walencyjnego oraz poziomu donorowego w odległości 0,36 eV od dna pasma przewodnictwa, a bombardowanie za pomocą neutronów — pojawienie się poziomu akceptorowego dokładnie w środku przerwy energetycznej. Wytworzone w ten sposób centra rekombinacyjne mogą być usunięte przez odprężanie półprzewodnika w podwyższonej temperaturze.

9.5* REKOMBINACJA AUGERA

W dotychczasowych rozważaniach nie uwzględnialiśmy wpływu zjawiska Augera na przebieg procesu rekombinacji nośników w półprzewodniku. Zjawisko to może występować zarówno przy rekombinacji bezpośredniej (w tym także w półprzewodnikach ze skośną przerwą energetyczną), jak i pośredniej, a energia wyzwolona w procesie rekombinacji może być przekazana albo elektronowi w pasmie przewodnictwa, albo też dziurze w pasmie walencyjnym. Wzbudzone w ten sposób nośniki tracą uzyskaną energię zazwyczaj wskutek zderzeń z siecią krystaliczną. Ostatecznym wynikiem całego procesu rekombinacji jest więc generacja pewnej liczby fononów.

Schematy różnych mechanizmów rekombinacji nośników w półprzewodniku zestawiono na rys. 9/9. W górnej części rysunku pokazano 3 możliwe mechanizmy rekombinacji bezpośredniej. Szybkość rekombinacji promienistej (rys. 9/9a), jak pamiętamy, jest proporcjonalna do koncentracji elektronów i do



9/9 Rekombinacja promienista (a) oraz różne rodzaje rekombinacji Augera: z przejściami bezpośrednimi (b, c) oraz z przejściami poprzez stany pułapkowe (d — g)

koncentracji dziur – zob. wzór (9/6). Szybkość rekombinacji bezpośredniej Augera jest ponadto proporcjonalna do koncentracji tych nośników, które przejmują energię wyzwalaną w procesie rekombinacji. Łącznie więc szybkość rekombinacji przedstawionej na rys. 9/9b jest proporcjonalna do iloczynu n^2p, a rekombinacji przedstawionej na rys. 9/9c – do iloczynu np^2. Oznaczając odpowiednie współczynniki proporcjonalności symbolami C_{An} i C_{Ap}, możemy szybkość rekombinacji Augera wyrazić wzorem

$$R_A = C_{An} n^2 p + C_{Ap} np^2 \tag{9/45}$$

Postępując analogicznie jak w p. 9.2, można wyznaczyć wypadkową szybkość rekombinacji U_{rg}. Przykładowo, dla półprzewodnika typu n przy małym poziomie zakłócenia wynosi ona

$$U_{rg} = C_{An} n_0^2 \Delta p \tag{9/46}$$

jest więc proporcjonalna do koncentracji nośników nadmiarowych. Z zależności tej wynika następujący wzór na czas życia związany z efektem Augera:

$$\tau_A = \frac{1}{C_{An} N_d^2} \tag{9/47}$$

Współczynnik C_{An} jest dla krzemu rzędu $1 \ldots 3 \cdot 10^{-31}$ cm^6/s. Wypływa stąd wniosek, że przy koncentracjach domieszek mniejszych niż 10^{18} cm^{-3}, rekombinacja Augera nie odgrywa istotnej roli. Czasy życia τ_A są w tych warunkach dłuższe niż typowe czasy życia, wynikające z rekombinacji poprzez centra rekombinacyjne. Dopiero przy koncentracjach domieszek rzędu 10^{19} cm^{-3} udział rekombinacji Augera zaczyna być istotny, a przy jeszcze większych koncentracjach rekombinacja ta staje się procesem dominującym.

9.6* REKOMBINACJA POWIERZCHNIOWA

Właściwości większości elementów półprzewodnikowych zależą nie tylko od zjawisk zachodzących we wnętrzu tworzących je monokryształów, lecz również od procesów rekombinacyjno-generacyjnych występujących na ich powierzchniach.

Na powierzchni kryształu urywa się raptownie regularna struktura sieci krystalicznej, a zatem załamuje się w sposób nagły periodyczność potencjału elektrycznego, na której opierały się wywody prowadzące do modelu pasmowego ciała stałego. W wyniku tego zakłócenia periodyczności sieci występują na powierzchni półprzewodnika dodatkowe stany elektronowe, nazywane *stanami powierzchniowymi*. Stanom tym odpowiadają poziomy energetyczne położone w przerwie energetycznej półprzewodnika.

Drugą przyczyną powstawania stanów powierzchniowych są defekty strukturalne powstające przy obróbce mechanicznej kryształu (cięcie, szlifowanie itd.), a także zanieszyszczenia w postaci obcych atomów i cząsteczek osadzających się na powierzchni kryształu.

W wyniku gwałtownego urwania się struktury krystalicznej, atomy znajdujące się w warstwie przypowierzchniowej kryształu mają część wiązań międzyatomowych niewykorzystanych. Wskutek istnienia takich wiązań, zwanych *wiązaniami niewysyconymi*, powierzchnia kryształu odznacza się szczególnie dużą aktywnością, tzn. łatwo przyłącza (adsorbuje) obce atomy i cząsteczki. Są one utrzymywane na powierzchni kryształu siłami elektrycznymi i zachowują swe cechy indywidualne (adsorpcja fizyczna) lub też tworzą połączenia chemiczne z materiałem podłoża (adsorpcja chemiczna).

Stany powierzchniowe występują w krysztale w warstwie przypowierzchniowej o grubości kilku stałych sieci. Ze względu na niewielką grubość tej warstwy, dogodnie jest uważać, że są one rozłożone na powierzchni kryształu, i operować ich gęstością powierzchniową, którą często oznacza się symbolem N_{ss}[1].

Gęstość stanów powierzchniowych na czystej powierzchni krzemu jest rzędu 10^{15} cm^{-2}. W większości elementów półprzewodnikowych powierzchnia ta jest pokryta dość grubą amorficzną warstwą dwutlenku krzemu SiO_2. Warstwa ta spełnia w tych elementach kilka ważnych funkcji, m.in. służy ona jako tzw. *warstwa pasywująca*, tzn. zabezpieczająca powierzchnię krzemu od wpływu czynników zewnętrznych.

Warstwę SiO_2 wytwarza się zazwyczaj przez utlenienie podłoża krzemowego, dzięki czemu powierzchnia graniczna układu krzem–dwutlenek krzemu powstaje w głębi pierwotnego kryształu. Taki sposób pasywacji powierzchni zapewnia jej dużą czystość. Ponadto obecność atomów tlenu w pobliżu powierzchni krzemu znacznie zmniejsza liczbę niewysyconych wiązań atomów tworzących monokryształ krzemu. W wyniku tego gęstość stanów powierzchniowych na tak otrzymanej powierzchni krzemu jest wiele tysięcy razy mniejsza niż podana poprzednio wartość. W zależności od orientacji krystalograficznej rozpatrywanej płaszczyzny, sposobu wytwarzania warstwy SiO_2 oraz jej dalszej obróbki technologicznej, gęstość stanów powierzchniowych zawiera się zwykle w granicach od 10^{10} do $5 \cdot 10^{11}$ cm^{-2}. Ostatnio opracowano technologię utleniania krzemu dającą wartości N_{ss} leżące poniżej granicy wykrywalności przy użyciu najdoskonalszych metod pomiarowych.

Kinetyka procesów rekombinacyjno-generacyjnych na powierzchni półprzewodnika jest podobna do analogicznych procesów zachodzących w jego wnętrzu i opisanych w p. 9.3. Podobnie jak w przypadku omawianych tam stanów pułapkowych, także na stanach powierzchniowych może następować lokalizacja elektronów i dziur. Stany te mogą więc być elektrycznie obojętne, albo też wykazywać dodatni lub ujemny ładunek elektryczny. W wyniku tego potencjał powierzchni półprzewodnika odbiega na ogół do potencjału panującego w jego wnętrzu.

Miarą intensywności procesów rekombinacyjno-generacyjnych we wnętrzu półprzewodnika jest szybkość rekombinacji zdefiniowana w p. 9.2. Intensywność rekombinacji (generacji) zachodzącej na powierzchni półprzewodnika wyrażamy

---

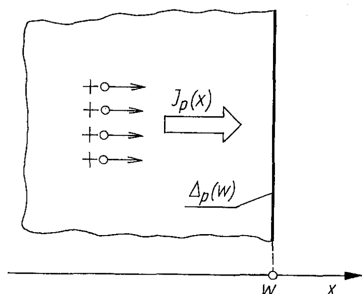
[1] Indeks *ss* pochodzi od angielskiego terminu „surface states" (stany powierzchniowe)

za pomocą innego parametru, który nazywamy *prędkością rekombinacji powierzchniowej*.

W celu zdefiniowania tej wielkości, rozpatrzmy sytuację przedstawioną na rys. 9/10. Załóżmy, że do powierzchni półprzewodnika dopływa strumień dziur, któremu odpowiada prąd o gęstości $J_p(x)$. Dochodząc do powierzchni półprzewodnika (tzn. do płaszczyzny $x = w$), dziury te ulegają rekombinacji, przy czym w pobliżu powierzchni ustala się nadmiarowa koncentracja dziur Δp. W stanie ustalonym liczba dziur dopływających do powierzchni jest oczywiście równa liczbie dziur ulegających rekombinacji na powierzchni. Warunek ten można wyrazić równaniem

$$J_p(w) = qs\,\Delta p\,(w) \qquad (9/48)$$

przy czym wprowadzonym tu współczynnikiem proporcjonalności s jest prędkość rekombinacji powierzchniowej.



9/10
Wielkości występujące w definicji prędkości rekombinacji powierzchniowej

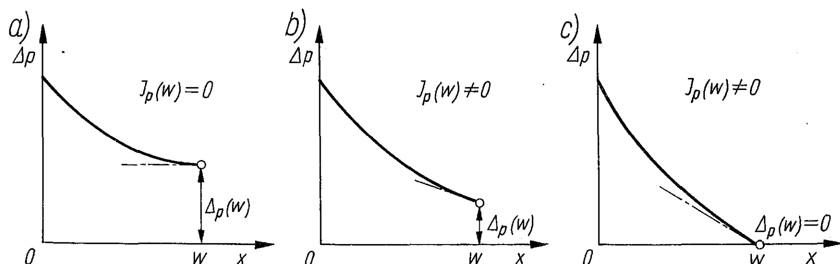
W przypadku granicznym, gdy $J_p(w) = 0$, a $\Delta p\,(w)$ ma wartość różną od zera, prędkość rekombinacji powierzchniowej jest równa zeru. Koncentracja nadmiarowa dziur utrzymuje się więc w tych warunkach na powierzchni bez dopływu dziur z wnętrza półprzewodnika. Przypadek ten pokazano na rys. 9/11a. Założono przy tym, że przepływ dziur w rozpatrywanym półprzewodniku następuje w wyniku ich dyfuzji w kierunku zgodnym z kierunkiem osi x, a gęstość wytworzonego przez nie prądu jest proporcjonalna do gradientu koncentracji dziur w danym punkcie[1)].

Drugi graniczny przypadek, przedstawiony na rys. 9/11c, odpowiada sytuacji gdy $\Delta p\,(w) = 0$ przy różnej od zera wartości prądu $J_p(w)$. Oznacza to, że prędkość rekombinacji powierzchniowej jest nieskończenie wielka. Taką praktycznie nieskończenie wielką prędkością rekombinacji powierzchniowej odznaczają się dobre kontakty metal-półprzewodnik.

Jeżeli prędkość rekombinacji powierzchniowej nie jest ani nieskończenie mała, ani też nieskończenie wielka, to $J_p(w) \neq 0$ oraz $\Delta p(w) \neq 0$. Sytuację taką pokazano na rys. 9/11b.

---

1) Zjawisko dyfuzji nośników ładunku i wytwarzanych przez nie prądów dyfuzyjnych będą omawiane w p. 10.7.



9/11 Rozkład koncentracji nadmiarowych dziur w półprzewodniku typu *n* przy różnych wartościach prędkości rekombinacji powierzchniowej: a) $s = 0$; b) $s \neq 0$; c) $s \to \infty$

Prędkość rekombinacji powierzchniowej zależy bardzo silnie od sposobu obróbki powierzchni półprzewodnika i od potencjału tej powierzchni w stosunku do wnętrza półprzewodnika. Dla powierzchni szlifowanych przekracza ona zwykle 10^4 cm/s, a po polerowaniu i trawieniu chemicznym może spaść, do wartości rzędu 10^2 cm/s. Dla powierzchni krzemu pokrytych warstwą pasywującą z SiO_2 można przy zachowaniu określonych warunków technologicznych uzyskać prędkości rekombinacji powierzchniowej mniejsze niż 1 cm/s.

10

PRZEPŁYW PRĄDU ELEKTRYCZNEGO W PÓŁPRZEWODNIKU

Działanie większości elementów półprzewodnikowych opiera się na wykorzystaniu prądów płynących przez odpowiednie końcówki danego elementu. Prądy te zależą od różnych czynników wewnętrznych (takich jak rozmiary poszczególnych obszarów elementu, rodzaj materiału, z którego jest on wykonany, koncentracje domieszek itd.) oraz zewnętrznych (napięcia doprowadzane do poszczególnych elektrod, promieniowanie zewnętrzne, pole magnetyczne itd.). W rozdziale tym zapoznamy się z różnymi rodzajami prądów, z jakimi spotykamy się w półprzewodnikach, oraz zależnościami tych prądów od niektórych z wymienionych wyżej czynników.

Niezależnie od uporządkowanych ruchów wywołanych czynnikami wymuszającymi, takimi np. jak pole elektryczne, nośniki ładunku w ciele stałym wykonują chaotyczne ruchy termiczne we wszystkich kierunkach. Ruchy te odbywają się zwykle ze znacznie większymi prędkościami niż ruchy uporządkowane, jednakże na skutek zupełnie przypadkowego rozkładu tych prędkości nie wywołują one wypadkowego przepływu nośników w żadnym kierunku. Ruchy termiczne nośników mają jednak istotne znaczenie przy rozpatrywaniu właściwości elektrycznych ciał stałych, dlatego też poświęcimy im odpowiednią uwagę.

10.1 PODSTAWOWE DEFINICJE

Prąd elektryczny jest to uporządkowany ruch ładunków elektrycznych. W metalach jest on związany z przepływem elektronów, w półprzewodnikach zaś — z przepływem 2 rodzajów nośników ładunku: elektronów i dziur. W ogólnym przypadku można więc gęstość prądu $\boldsymbol{J}$ wyrazić jako sumę gęstości prądu *elektronowego* $\boldsymbol{J}_n$ i *dziurowego* $\boldsymbol{J}_p$

$$\boldsymbol{J} = \boldsymbol{J}_n + \boldsymbol{J}_p \qquad (10/1)$$

Przepływ nośników ładunku w półprzewodniku może być wywołany działaniem różnych czynników, w zależności od tego nadajemy odpowiednim prądom różne nazwy. Rozróżniamy więc:

a) prąd wywołany polem elektrycznym. Siły tego pola zmuszają chaotycznie poruszające się nośniki do przesuwania się w kierunku zgodnym (dziury) lub przeciwnym (elektrony) do kierunku pola. Można więc powiedzieć, że nośniki ładunku w półprzewodniku są *unoszone* przez pole elektryczne w dwóch przeciwnych kierunkach. Z tego względu odpowiedni składnik prądu nazywamy *prądem unoszenia*;

b) prąd wywołany nierównomiernym rozkładem koncentracji nośników. W wyniku ruchów termicznych nośniki przemieszczają się z obszarów o większej do obszarów o mniejszej koncentracji. Proces ten prowadzi w granicy do wyrównania się koncentracji nośników w całej objętości półprzewodnika. Nosi on nazwę *dyfuzji*, a odpowiedni składnik prądu jest nazywany *prądem dyfuzyjnym*;

c) prąd wywołany nierównomiernym rozkładem temperatury w półprzewodniku. Temperatura jest miarą intensywności ruchów termicznych, którym podlegają nośniki w ciele stałym. W obszarach o wyższej temperaturze nośniki mają większe energie, a więc poruszają się z większymi prędkościami. Mają one zatem tendencję do przemieszczania się w kierunku obszarów o niższej temperaturze. Proces ten nosi nazwę *termodyfuzji*, a wywołany nim składnik prądu nazywamy *prądem termodyfuzyjnym*.

Podana na początku tego punktu definicja prądu elektrycznego nie jest w pełni poprawna, nie obejmuje bowiem wprowadzonego przez Maxwella pojęcia *prądu przesunięcia*[1]. Uwzględniając ten prąd, należy wzór (10/1) doprowadzić do postaci

$$\boldsymbol{J} = \boldsymbol{J}_n + \boldsymbol{J}_p + \boldsymbol{J}_D \tag{10/2}$$

Pierwsze dwa składniki łącznie stanowią gęstość prądu *przewodzenia*, trzeci składnik natomiast jest gęstością prądu przesunięcia.

Prąd przesunięcia jest zdefiniowany jako szybkość zmian strumienia elektrycznego przechodzącego przez rozpatrywaną powierzchnię. Gęstość tego prądu wyraża się więc wzorem

$$\boldsymbol{J}_D = \frac{\partial \boldsymbol{D}}{\partial t} \tag{10/3}$$

gdzie $\boldsymbol{D}$ oznacza indukcję elektryczną.

Ze wzoru tego wynika, że prąd przesunięcia występuje wówczas, gdy indukcja elektryczna zmienia się w czasie, a to zachodzi w następujących dwóch przypadkach:

a) gdy zmieniają się napięcia elektrod wytwarzających pole elektryczne w rozpatrywanym obszarze półprzewodnika (lub dielektryka);

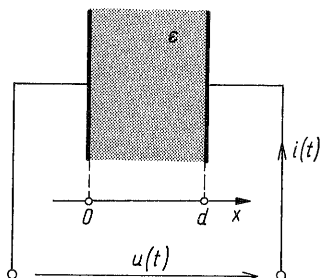
b) gdy rozkład ładunków elektrycznych w półprzewodniku ulega zmianie.

Najprostszy przypadek, w którym występuje prąd przesunięcia przedstawiono na rys. 10/1. Pokazano tam kondensator płaski wypełniony dielektrykiem o przenikalności elektrycznej ε. W rozważaniach naszych założymy, że w dielektryku nie występują żadne nośniki ładunku. Jeżeli do zacisków kondensatora doprowadzimy zmieniające się w czasie napięcia $u(t)$, to indukcja elektryczna w dowolnym przekroju x będzie równa

$$D(t) = \varepsilon E(t) = \varepsilon \frac{u(t)}{d} \tag{10/4}$$

---

[1] Ściślej można zdefiniować prąd elektryczny jako źródło wirowego pola magnetycznego. Pojęcie to obejmuje wówczas zarówno prąd przewodzenia związany z przepływem ładunków elektrycznych, jak i prąd przesunięcia.



10/1
Kondensator płaski

Gęstość występującego w dielektryku prądu przesunięcia wyniesie zatem

$$J(t) = \frac{\partial D}{\partial t} = \frac{\varepsilon}{d} \frac{\mathrm{d}u}{\mathrm{d}t} \tag{10/5}$$

a całkowity prąd przepływający przez kondensator będzie miał wartość

$$i(t) = \frac{\varepsilon A}{d} \frac{\mathrm{d}u}{\mathrm{d}t} \tag{10/6}$$

gdzie A – powierzchnia okładek kondensatora.

Wyrażenie $\varepsilon A/d$ reprezentuje pojemność kondensatora płaskiego. Uwzględniając to, można wzór (10/6) napisać w następującej ogólniejszej postaci słusznej dla dowolnego rodzaju kondensatora:

$$i(t) = C \frac{\mathrm{d}u}{\mathrm{d}t} \tag{10/7}$$

Prąd przesunięcia płynący w dielektryku pozbawionym nośników ładunku jest to więc znany nam *prąd pojemnościowy* związany z ładowaniem lub rozładowywaniem kondensatora w wyniku zmiany napięcia doprowadzonego do jego okładek.

Prąd całkowity, będący sumą prądu wywołanego przepływem ładunków elektrycznych i prądu przesunięcia, podlega *prawu ciągłości*, tzn. ma on jednakową wartość w dowolnym przekroju nierozgałęzionego obwodu elektrycznego. Prawo to nie dotyczy natomiast poszczególnych składników występujących we wzorze (10/2). W książce tej spotkamy się często z przypadkami, w których wartości tych składników będą zależne od współrzędnej mierzonej wzdłuż rozpatrywanej próbki półprzewodnika.

Dalsze rozważania zawarte w tym rozdziale będą ograniczone do tzw. *przepływów stacjonarnych*, tj. przepływów, które występują w stanie ustalonym przy stałych wartościach napięć doprowadzonych do elektrod. W tych warunkach indukcja elektryczna w dowolnym punkcie nie zmienia się w czasie, a zatem prąd przesunięcia nie występuje. Prąd elektryczny jest wówczas związany jedynie z przepływem nośników ładunku.

10.2 RUCHY TERMICZNE NOŚNIKÓW W PÓŁPRZEWODNIKU

Jak już mówiliśmy w p. 4.6, atomy w krysztale wykonują drgania wokół swych położeń równowagi. Energia tych drgań udziela się również elektronom, które w związku z tym poruszają się ze znacznymi prędkościami i zderzają z siecią krystaliczną wskutek jej niedoskonałości (odchylenia atomów od położeń równowagowych, defekty strukturalne, domieszki itd.). Rozkład kierunkowy prędkości ruchów termicznych jest izotropowy (tzn. każdy kierunek prędkości jest jednakowo prawdopodobny), a rozkład ze względu na wartości tych prędkości jest opisany funkcją Maxwella-Boltzmanna.

Wartość średnia prędkości ruchów termicznych jest równa zeru, tzn.

$$\langle v \rangle = \frac{1}{N} \sum_i v_i = 0 \tag{10/8}$$

gdzie N – liczba rozpatrywanych elektronów.

Wniosek ten wynika stąd, że w warunkach równowagi termodynamicznej przepływ nośników przez dowolną płaszczyznę w jednym kierunku jest równoważony (w sensie statystycznym) przez taki sam przepływ w kierunku przeciwnym.

Nie oznacza to oczywiście, że energia tych ruchów jest także równa zeru. Z praw fizyki statystycznej wynika bowiem, że dla każdej cząstki układu znajdującego się w równowadze termodynamicznej na każdy stopień swobody przypada średnio energia równa $\frac{1}{2}kT$. Ponieważ elektron ma 3 stopnie swobody, zatem jego średnia energia kinetyczna w ciele stałym wynosi $\frac{3}{2}kT$, co można wyrazić wzorem

$$\frac{1}{2} m_n^* v_{th}^2 = \frac{3}{2} kT \tag{10/9}$$

gdzie m_n^* – masa efektywna, a v_{th} – tzw. *prędkość termiczna* elektronów[1]. We wzorze tym występuje temperatura kryształu T, w rozpatrywanych warunkach elektrony są bowiem w równowadze cieplnej z siecią krystaliczną.

Ze wzoru (10/9) otrzymujemy

$$v_{th} = \sqrt{\frac{3kT}{m_n^*}} \tag{10/10}$$

Podstawiając $T = 300$ K i przyjmując dla prostoty $m_n^* = m_0$, znajdziemy, że w temperaturze pokojowej prędkość termiczna elektronów w ciele stałym jest rzędu 10^7 cm/s.

Analogiczne wzory można napisać dla dziur; należy w nich jedynie zamiast m_n^* podstawić masę efektywną m_p^*.

Drugą ważną wielkością związaną z ruchem elektronów w ciele stałym jest *średni czas między kolejnymi zderzeniami* elektronu z siecią krystaliczną. Zależy on od intensywności drgań termicznych sieci (a więc od jej temperatury) i od koncentracji defektów strukturalnych. Dla nieograniczonej doskonale periodycznej sieci czas ten byłby nieskończenie wielki. Na podstawie rozważań z rozdz. 6, dotyczących propagacji fali elektronowej w krysztale, wiemy bowiem, że elektron o energii zawierającej się w dozwolonym pasmie energetycznym porusza się bez żadnych zakłóceń ze strony sieci krystalicznej, tzn. nie ulega zderzeniom z tą siecią. Oddziaływanie idealnie periodycznej sieci na ruch elektronu uwzględnia się, przypisując mu pewną masę efektywną różniącą się od masy elektronu swobodnego w próżni.

[1] Jest to wartość średnia kwadratowa prędkości, zdefiniowana wzorem

$$v_{th} = \sqrt{\frac{1}{N} \sum_i v_i^2}$$

W kryształach wykorzystywanych w elementach półprzewodnikowych średni czas między kolejnymi zderzeniami elektronu z siecią krystaliczną zawiera się w granicach $10^{-13}-10^{-11}$ s.

Znając ten czas oraz prędkość termiczną elektronów, można wyznaczyć tzw. *średnią drogę swobodną* elektronu, czyli średnią odległość przebywaną przez elektron między kolejnymi zderzeniami z siecią. W interesujących nas półprzewodnikach wynosi ona zwykle $10^{-8}-10^{-6}$ m. Ponieważ stała sieci tych kryształów wynosi kilka dziesiętnych nanometra, zatem elektron przebywa między kolejnymi zderzeniami odległość zawartą w granicach od kilkudziesięciu do kilku tysięcy stałych sieci[1].

Pomimo tego, że nośniki ładunku w półprzewodniku poruszają się we wszystkich kierunkach ze stosunkowo dużymi prędkościami, nie wytwarzają one uśrednionego w czasie prądu w żadnym kierunku. Istnieje jedynie mierzalna składowa zmienna prądu związana z przypadkowym charakterem ruchów elektronów. Objawia się ona w postaci tzw. *fluktuacji termicznych* i jest odpowiedzialna za tzw. *szumy „cieplne”* występujące w każdym przewodniku. Szumy te mają w elektronice istotne znaczenie, ponieważ są one jednym z czynników określających próg czułości elementów przeznaczonych do wzmacniania słabych sygnałów elektrycznych.

10.3 PRĄD UNOSZENIA

Rozpatrzmy teraz wpływ pola elektrycznego na zachowanie się nośników ładunku w półprzewodniku. Weźmy najpierw pod uwagę elektrony w pasmie przewodnictwa. W nieobecności pola poruszają się one, jak to już powiedziano, z przypadkowymi prędkościami we wszystkich kierunkach. Przedstawiono to w sposób poglądowy na rys. 10/2a.

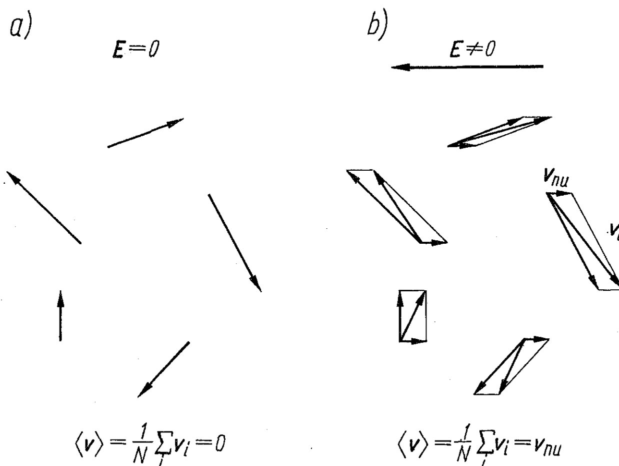
Gdy w półprzewodniku wytworzymy pole elektryczne o natężeniu E, to na wszystkie elektrony będzie wywierana siła w tym samym kierunku i na nieuporządkowany ruch elektronów nałoży się dodatkowy, uporządkowany ruch w kierunku przeciwnym do kierunku pola.

Gdyby ruch elektronów odbywał się bez zakłóceń, tak jak to występuje na przykład w próżni, to ta dodatkowa składowa ruchu wywołana oddziaływaniem pola elektrycznego byłaby ruchem jednostajnie przyspieszonym wzdłuż linii tego pola. W ciele stałym elektrony doznają jednakże ciągłych zderzeń z zaburzeniami sieci krystalicznej, przy czym przy każdym takim zderzeniu następuje raptowna i całkowicie przypadkowa zmiana wartości i kierunku prędkości elektronu.

Na tle tych chaotycznych ruchów uwidocznia się jednakże stała tendencja elektronów do przesuwania się w kierunku działania sił pola elektrycznego.

---

[1] Otrzymano jednakże kryształy, w których średnia droga swobodna elektronów w temperaturze pokojowej wynosi kilka cm, jest więc setki milionów razy większa od stałej sieci.



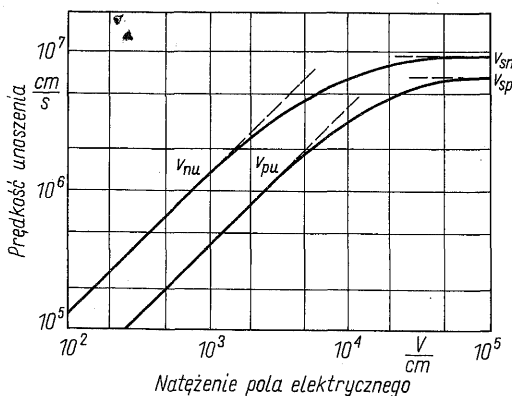
10/2 Rozkład prędkości elektronów w warunkach równowagi termodynamicznej (a) i w obecności pola elektrycznego (b)

Każdemu elektronowi można zatem przypisać pewną dodatkową składową prędkości skierowaną w kierunku przeciwnym do kierunku wektora $\boldsymbol{E}$. Uwidoczniono to na rys. 10/2b.

Średnia prędkość elektronów przyjmuje teraz różną od zera wartość, którą nazywamy *prędkością unoszenia* i którą oznaczamy symbolem $\boldsymbol{v}_{nu}$. Jest ona zdefiniowana wzorem

$$\boldsymbol{v}_{nu} = \frac{1}{N}\sum_i \boldsymbol{v}_i \qquad (10/11)$$

Jak wynika z przeprowadzonych doświadczeń, prędkość v_{nu} zależy od natężenia pola elektrycznego w sposób przedstawiony przykładowo na rys. 10/3. Wykresy pokazane na tym rysunku dotyczą krzemu. Podobny charakter mają one dla wielu innych półprzewodników. Istnieją jednakże materiały, do których należy na przykład GaAs, w których prędkość unoszenia nie jest mo-



10/3
Zależność prędkości unoszenia elektronów i dziur w krzemie w temperaturze 300 K od natężenia pola elektrycznego

notonicznie wzrastającą funkcją natężenia pola elektrycznego. Zagadnienia te będą omawiane w p. 10.6.2.

Z rys. 10/3 wynika, że w zakresie niezbyt silnych pól elektrycznych (np. dla krzemu przy $E \leqslant 10^3$ V/cm) prędkość unoszenia elektronów (a podobny wniosek dotyczy także dziur) jest proporcjonalna do natężenia pola elektrycznego. Wynika to nie tylko z prostoliniowości odpowiedniego odcinka charakterystyki przedstawionej na rys. 10/3, lecz również z wartości współczynnika kierunkowego tej prostej. Uwzględniając to, że prędkość v_{nu} jest skierowana przeciwnie niż wektor E, można zatem dla zakresu proporcjonalności napisać

$$v_{nu} = -\mu_n E \tag{10/12}$$

przy czym wprowadzony tu współczynnik proporcjonalności μ_n nosi nazwę *ruchliwości elektronów.*

Zależność (10/12) można rozszerzyć również na nieliniowy odcinek charakterystyki $v_{nu}(E)$, wówczas jednak współczynnik μ_n nie będzie wielkością stałą, lecz będzie zależny od natężenia pola elektrycznego.

W miarę powiększania natężenia pola przyrost prędkości unoszenia staje się coraz wolniejszy, a po przekroczeniu pewnej wartości (która dla krzemu wynosi ok. $5 \cdot 10^4$ V/cm) prędkość unoszenia staje się niezależna od E. Mówimy wówczas, że prędkość unoszenia uległa *nasyceniu.* Wartość prędkości w tym zakresie nazywamy *prędkością nasycenia* i oznaczamy symbolem v_{sn}.

Prędkość unoszenia dziur v_{pn} jest skierowana zgodnie z kierunkiem wektora E, a zatem odpowiednią zależność dla dziur można wyrazić wzorem

$$v_{pu} = \mu_p E \tag{10/13}$$

przy czym μ_p oznacza *ruchliwość dziur.*

Na rysunku 10/3 jest widoczne, że przy dowolnej wartości natężenia pola elektrycznego prędkość unoszenia jest dla dziur mniejsza niż dla elektronów — dotyczy to nie tylko krzemu, lecz również większości innych półprzewodników. Z faktu tego wynika ważny wniosek, że elementy półprzewodnikowe, których działanie opiera się na wykorzystaniu przepływu elektronów (np. tranzystory *npn*) odznaczają się większą szybkością działania niż analogiczne elementy, w których jest wykorzystywany przepływ dziur (np. tranzystory *pnp*).

W tablicy 10/1 zestawiono niektóre dane dotyczące prędkości unoszenia nośników dla kilku półprzewodników samoistnych. Z zestawienia tego wynika, że w zależności od rodzaju materiału i rodzaju nośnika ruchliwości mogą różnić się o kilka rzędów wielkości. Jak zobaczymy w następnym punkcie, inne czynniki (a zwłaszcza wprowadzenie domieszek) mają również duży wpływ na wartość ruchliwości nośników.

Prędkości nasycenia różnią się natomiast stosunkowo nieznacznie i są tego samego rzędu co prędkości termiczne nośników w półprzewodniku.

Pole elektryczne może być wytworzone w półprzewodniku albo przez źródła zewnętrzne, albo też przez ładunki elektryczne znajdujące się w rozpatrywanej próbce półprzewodnika lub na jej powierzchni. Przeanalizujemy ten pierwszy przypadek, tzn. rozpatrzmy jednorodny półprzewodnik umieszczony

10/1 Wartości parametrów charakteryzujących zależność prędkości nośników ładunku od natężenia pola elektrycznego dla półprzewodników samoistnych ($T = 300$ K)

Półprzewodnik	Rodzaj nośników	Ruchliwość przy $E < E_1$	E_1	E_2	Prędkość nasycenia
		$cm^2/V \cdot s$	V/cm	V/cm	cm/s
Si	elektrony	1415	10^3	$5 \cdot 10^4$	10^7
	dziury	470	$3 \cdot 10^3$	10^5	$6{,}5 \cdot 10^6$
Ge	elektrony	3900	$8 \cdot 10^2$	$5 \cdot 10^3$	$6 \cdot 10^6$
	dziury	1900	$2 \cdot 10^2$		
GaAs	elektrony	8500	$2 \cdot 10^3$	$3 \cdot 10^4$	$8 \cdot 10^6$
	dziury	400			

Objaśnienia: E_1 – natężenie pola, do którego rozciąga się zakres proporcjonalności;
E_2 – natężenie pola, od którego rozpoczyna się zakres nasycenia.

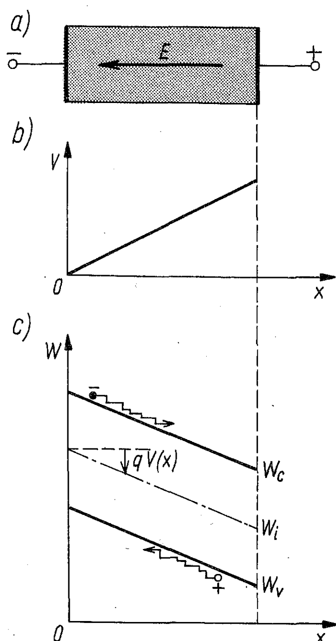
w polu elektrycznym. W celu uproszczenia rozważań przyjmijmy, że mamy do czynienia z prostopadłościenną próbką przedstawioną na rys. 10/4a. Próbka ta jest zakończona na dwóch przeciwległych ściankach idealnymi kontaktami, za pomocą których jest ona włączona do obwodu elektrycznego[1]. Do kontaktów tych doprowadzone jest stałe napięcie o polaryzacji pokazanej na rysunku. W wyniku tego w próbce powstaje jednorodne pole elektryczne o natężeniu E.

Rozkład potencjału w próbce przedstawiono na rys. 10/4b. Znając ten rozkład, można wykreślić model pasmowy rozpatrywanego półprzewodnika – rys. 10/4c. W modelu tym na osi rzędnych jest odkładana, jak wiadomo, energia elektronu. Ponieważ między energią potencjalną elektronu a potencjałem elektrycznym zachodzi związek

$$W(x) = -q\,V(x) \tag{10/14}$$

zatem linie odpowiadające energiom W_c, W_i oraz W_v będą nachylone w stosunku do osi x pod kątem zależnym od natężenia pola elektrycznego w półprze-

[1] *Idealny kontakt* charakteryzuje się tym, że przy przepływie prądu w dowolnym kierunku nie występuje na nim spadek napięcia, a koncentracje nośników w bezpośrednim jego sąsiedztwie są równe koncentracjom równowagowym, niezależnie od tego, co dzieje się wewnątrz próbki półprzewodnika. Ta ostatnia cecha oznacza, że idealny kontakt stanowi powierzchnię graniczną o nieskończenie wielkiej prędkości rekombinacji powierzchniowej. Z tego względu kontakt taki nazywamy często kontaktem *rekombinacyjnym*. Dobrze wykonane kontakty rzeczywiste można w większości przypadków uznać za spełniające w znacznym stopniu te warunki.



10/4
Zachowanie się nośników ładunku pod działaniem pola elektrycznego w półprzewodniku: a) próbka półprzewodnika; b) rozkład potencjału; c) model pasmowy półprzewodnika

wodniku, zgodnie ze wzorem

$$E = -\frac{dV}{dx} = \frac{1}{q}\frac{dW_i}{dx} \tag{10/15}$$

Na elektrony w pasmie przewodnictwa działa siła skierowana przeciwnie do kierunku wektora **E**. Elektrony te poruszają się więc w kierunku kontaktu o potencjale dodatnim. Na modelu pasmowym jest widoczne, jak elektron w pasmie przewodnictwa przesuwa się wzdłuż osi x wskutek pochylenia dna tego pasma. Pod pewnym względem ruch tego elektronu przypomina zsuwanie się punktu materialnego po równi pochyłej. Istnieje jednakże istotna różnica między tymi dwoma ruchami. Punkt materialny, zsuwając się bez tarcia po równi pochyłej zachowuje stałą wartość energii. Jego energia potencjalna maleje wprawdzie stopniowo, jednakże jej kosztem wzrasta energia kinetyczna. Elektron w półprzewodniku „zsuwając się" po dnie pasma przewodnictwa traci natomiast nieustannie energię wskutek zderzeń z niedoskonałościami sieci krystalicznej. Pokazano to w sposób poglądowy na rys. 10/4c. Linia obrazująca zmiany energii elektronu jest linią schodkową, przy czym długości poszczególnych schodków odpowiadają odległościom przebywanym przez elektron między kolejnymi zderzeniami, a wysokości schodków – porcjom energii przekazywanym sieci w kolejnych zderzeniach.

Przy przemieszczaniu się ładunków elektrycznych pod działaniem sił pola elektrycznego następuje, jak wiadomo, pobór energii z tego pola. W rozpatrywanym przypadku energia ta jest pobierana ze źródła zasilającego, a wydziela się w sieci krystalicznej, powodując wzrost intensywności drgań atomów, tzn.

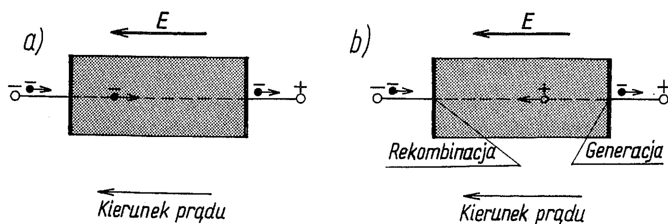
powiększając temperaturę kryształu. Zjawisko to występuje przy przepływie prądu elektrycznego w każdym przewodniku, przy czym energia wydzielana przez prąd jest nazywana *ciepłem Joule'a.*

Dziury w półprzewodniku zachowują się analogicznie jak elektrony, z tym, że jako nośniki obdarzone ładunkiem dodatnim, poruszają się one w przeciwnym kierunku niż elektrony. Podobnie też na modelu pasmowym energia dziur jest mierzona w przeciwnym kierunku, a zatem dziura tracąc energię w zderzeniach z siecią krystaliczną przesuwa się na rysunku w górę wzdłuż linii stanowiącej wierzchołek pasma walencyjnego – zob. rys. 10/4c. Jeżeli ruch elektronu w pasmie przewodnictwa przyrównaliśmy w pewnym sensie do zsuwania się punktu materialnego po równi pochyłej, to ruch dziury można porównać z zachowaniem się pęcherzyka powietrza w naczyniu wypełnionym cieczą (np. w poziomicy).

Rozpatrując przepływ elektronów i dziur przez próbkę półprzewodnika przedstawioną na rys. 10/4a, należy zwrócić uwagę na jeszcze jedną różnicę dotyczącą tych dwóch rodzajów nośników ładunku. Jak pokazano na rys. 10/5a, elektrony przepływające przez próbkę są wprowadzane i wyprowadzane z niej poprzez kontakty łączące tę próbkę z zewnętrznym obwodem elektrycznym. W przypadku dziur sytuacja jest nieco inna, ponieważ te quasi-cząstki są tworem fikcyjnym, reprezentującym zachowanie się wszystkich elektronów w pasmie walencyjnym. Dziura nie może więc „wyjść" z półprzewodnika poprzez kontakt do obwodu zewnętrznego. Zgodnie z rys. 10/5b dziury pojawiają się w półprzewodniku w wyniku generacji par elektron-dziura występującej na powierzchni jednego kontaktu, a znikają wskutek rekombinacji na powierzchni drugiego kontaktu. W obwodzie zewnętrznym poruszają się natomiast elektrony, które dopływają do kontaktu „rekombinacyjnego", a odpływają od kontaktu „generacyjnego".

Pomimo tego, że dziury poruszają się w półprzewodniku w przeciwnym kierunku niż elektrony, prądy wytwarzane przez oba rodzaje nośników płyną w tym samym kierunku – zgodnym z kierunkiem pola elektrycznego. Znając prędkości unoszenia nośników możemy przystąpić do wyznaczenia natężenia tych prądów.

Ograniczmy się początkowo do tzw. *przepływu jednowymiarowego*, tj. takiego, w którym wszystkie wektory są równoległe do osi, którą przyjmiemy



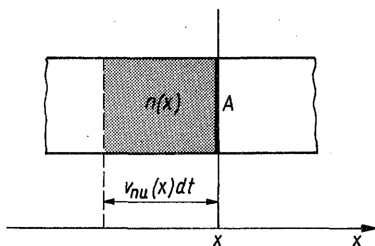
10/5 Prąd wytwarzany w półprzewodniku wskutek przepływu elektronów (a) i dziur (b)

za oś x, a wszystkie wielkości zależą tylko od tej jednej współrzędnej przestrzennej (i ewentualnie od czasu).

W celu wyznaczenia gęstości prądu elektronowego w danym przekroju x wyznaczmy ładunek przepływający w czasie dt przez powierzchnię A leżącą w płaszczyźnie x = const – rys. 10/6. Jeżeli prędkość unoszenia elektronów w punkcie x wynosi $v_{nu}(x)$, to w czasie dt przepływa przez rozpatrywaną powierzchnię ładunek zawarty w objętości $Av_{nu}\,dt$. Oznaczając koncentrację elektronów w tym obszarze symbolem $n(x)$, znajdziemy, że prąd elektronowy przepływający przez powierzchnię A wynosi $-Aqn(x)\,v_{nu}(x)$. Gęstość tego prądu ma więc wartość

$$J_{nu}(x) = -qn(x)\,v_{nu}(x) \tag{10/16}$$

przy czym zaznaczone wielkości są, ogólnie biorąc, funkcjami współrzędnej x.



10/6
Wielkości występujące w definicji gęstości prądu unoszenia

Ogólniejszy wzór, słuszny również dla przepływu wielowymiarowego, ma postać

$$\boldsymbol{J}_{nu} = -qn\boldsymbol{v}_{nu} \tag{10/17}$$

W podobny sposób otrzymuje się następujące wyrażenie na gęstość prądu dziurowego:

$$\boldsymbol{J}_{pu} = qp\boldsymbol{v}_{pu} \tag{10/18}$$

Korzystając z zależności (10/12) i (10/13), można wzory te przedstawić w postaci

$$\boxed{\begin{aligned} \boldsymbol{J}_{nu} &= qn\mu_n \boldsymbol{E} \quad (10/19)\\ \boldsymbol{J}_{pu} &= qp\mu_p \boldsymbol{E} \quad (10/20)\end{aligned}}$$

Gęstość całkowitego prądu unoszenia wynosi zatem

$$\boldsymbol{J}_u = \boldsymbol{J}_{nu} + \boldsymbol{J}_{pu} = q(n\mu_n + p\mu_p)\,\boldsymbol{E} \tag{10/21}$$

Prąd unoszenia ma więc kierunek zgodny z kierunkiem wektora $\boldsymbol{E}$ i w zakresie niezbyt silnych pól elektrycznych (tzn. w zakresie, w którym można uważać, że koncentracje nośników i ich ruchliwości są niezależne od E) jest on proporcjonalny do natężenia tego pola. Stwierdzenie to jest treścią *prawa Ohma*, które w zapisie wektorowym może być wyrażone wzorem

$$\boxed{\boldsymbol{J}_u = \sigma \boldsymbol{E}} \tag{10/22}$$

Występujący w tym równaniu współczynnik σ nosi nazwę *przewodności elektrycznej właściwej* lub *konduktywności*, a jego odwrotność, oznaczana symbolem ϱ – *oporu elektrycznego właściwego* lub *rezystywności.*

Na podstawie wzoru (10/21) można stwierdzić, że przewodność właściwa półprzewodnika wynosi

$$\boxed{\sigma = q\,(n\mu_n + p\mu_p)} \tag{10/23}$$

Jest ona, jak widzimy, zależna od koncentracji nośników i od ich ruchliwości, zależy więc nie tylko od własności danego materiału w określonej temperaturze, lecz również od warunków, w których znajduje się rozpatrywany obszar. Przykładowo, wprowadzając do danego obszaru nośniki z zewnątrz (np. wstrzykując je z przylegającego obszaru) zwiększamy ich koncentrację i wskutek tego powiększamy przewodność właściwą półprzewodnika. Podobny efekt wywołuje oświetlenie półprzewodnika promieniowaniem o dostatecznie dużej energii fotonów. W obu tych przypadkach mówimy o *modulacji* przewodności półprzewodnika.

Zwykle jednak, mówiąc o przewodności danego półprzewodnika jednorodnego, mamy na myśli wartość tego parametru w warunkach, w których koncentracje nośników są równe koncentracjom równowagowym.

Problem 10.1. Jaką postać przybiera wzór na przewodność właściwą półprzewodnika samoistnego (wielkość tę nazywamy *przewodnością samoistną* i oznaczamy symbolem σ_i)? Wyznacz przewodność właściwą, opór właściwy oraz współczynnik temperaturowy przewodności dla samoistnego krzemu w temperaturze pokojowej, zakładając, że ruchliwości nośników w tym półprzewodniku są proporcjonalne do $T^{-2,3}$. Przedyskutuj uzyskane wyniki.

Wzór (10/22) jest oczywiście słuszny także dla metali. W metalach nie ma jednakże konieczności zaznaczania, że zależność ta dotyczy prądu unoszenia, ponieważ w metalach, w przeciwieństwie do półprzewodników, nie występują inne prądy.

Podobnie, wzór (10/23) może być wykorzystywany dla metali, z tym zastrzeżeniem, że w metalach poza pasmami całkowicie obsadzonymi występuje tylko jedno pasmo częściowo obsadzone przez elektrony. Z tego powodu występuje w metalach tylko jeden rodzaj nośników ładunku – elektrony[1].

Problem 10.2. Przewodność właściwa miedzi w temperaturze pokojowej wynosi $5{,}88\cdot10^5\ (\Omega\cdot\text{cm})^{-1}$, a koncentracja elektronów w tym metalu ok. $10^{23}\ \text{cm}^{-3}$. Wyznacz ruchliwość elektronów w miedzi. Ile czasu zużywa elektron na przebycie drogi 1 m wzdłuż przewodnika miedzianego przy prądzie o gęstości 6 A/mm^2 (jest to dopuszczalna gęstość prądu w przewodach sieciowych urządzeń elektronicznych)? Jaką wartość ma w tych warunkach natężenie pola elektrycznego? Przedyskutuj uzyskane wyniki.

---

[1] Równie dobrze można przyjąć, że nośnikami tymi są dziury. Jest to szczególnie uzasadnione w odniesieniu do metali, w których najwyższe pasmo energetyczne zawiera mniej stanów nie obsadzonych przez elektrony niż stanów obsadzonych. Metale te, podobnie jak półprzewodniki typu p, charakteryzują się dodatnim znakiem stałej Halla i dodatnią siłą termoelektryczną. Zagadnienia te wykraczają jednakże poza ramy tej książki.

10.4 RUCHLIWOŚĆ NOŚNIKÓW

W poprzednim punkcie podane były wzory (10/12) i (10/13) wyrażające zależność prędkości unoszenia nośników od natężenia pola elektrycznego. Wzory te opisują zjawisko przepływu nośników pod działaniem pola elektrycznego od strony *makroskopowej*, tzn. zbiór nośników danego rodzaju jest traktowany całościowo, a uzyskana zależność wynika z pomiaru czasu przelotu nośników przez obszar o określonej długości lub z pomiaru prądu wytwarzanego przez te nośniki. Obecnie spróbujemy spojrzeć na zagadnienie w sposób *mikroskopowy*, rozpatrując zachowanie pojedynczych nośników. Oprzemy się przy tym na najprostszym możliwym modelu opisującym przepływ elektronów i dziur w ciele stałym.

Punktem wyjścia naszych rozważań będzie równanie ruchu elektronu w polu elektrycznym. Wzór (1/2), który omawialiśmy w p. 1.1, dotyczył ruchu elektronu nie podlegającego żadnym dodatkowym oddziaływaniom, a więc np. ruchu elektronu w próżni. Rozpatrując ruch elektronu w ciele stałym, musimy w równaniu tym uwzględnić zderzenia elektronu z siecią krystaliczną. Działają one hamująco na ruch elektronów, a więc odgrywają rolę podobną do sił tarcia jakiemu podlega np. kulka poruszająca się w ośrodku płynnym (tzn. w cieczy lub gazie). Ośrodek taki charakteryzuje się pewną lepkością, dlatego też tarcie to nazywamy *tarciem lepkim.*

Opierając się na powyższej analogii, przyjmiemy, że podobnie jak w przypadku tarcia lepkiego, siła przeciwdziałająca przemieszczaniu się elektronów w ciele stałym jest proporcjonalna do prędkości unoszenia elektronów[1]. Zgodnie z tym założeniem można równanie ruchu elektronu w ciele stałym przedstawić w następującej postaci:

$$\frac{\mathrm{d}v_{nu}}{\mathrm{d}t} = -\frac{q}{m_n^*}E - \frac{v_{nu}}{\tau} \tag{10/24}$$

gdzie τ jest odpowiednio dobranym współczynnikiem proporcjonalności, który dla późniejszej wygody umieszczony został w mianowniku.

W celu znalezienia interpretacji fizycznej tego współczynnika, rozważmy jak zmieniałaby się prędkość unoszenia elektronów, gdyby nagle usunąć pole elektryczne, tzn. w chwili $t = 0$ uczynić $E = 0$. Równanie (10/24) zredukowałoby się wówczas do postaci

$$\frac{\mathrm{d}v_{nu}}{\mathrm{d}t} = -\frac{v_{nu}}{\tau} \tag{10/25}$$

---

[1] W hydromechanice wzór określający wartość siły oporu, działającej na ciało kuliste poruszające się powoli w lepkim płynie nosi nazwę *wzoru Stokesa.* Zgodnie z tym wzorem siła ta jest proporcjonalna do prędkości ruchu ciała. Sile takiej podlegają na przykład krople deszczu spadające w powietrzu pod wpływem siły ciążenia. W pobliżu powierzchni ziemi poruszają się one ruchem jednostajnym z prędkością proporcjonalną do przyspieszenia ziemskiego g. Podobnie elektrony w ciele stałym poruszają się pod wpływem pola elektrycznego ze stałą prędkością unoszenia proporcjonalną do natężenia tego pola.

Rozwiązaniem tego równania jest funkcja

$$v_{nu}(t) = v_{nu}(0) \exp\left(-\frac{t}{\tau}\right) \tag{10/26}$$

Po usunięciu pola elektrycznego prędkość unoszenia elektronów, a wraz z nią również prąd unoszenia, zanikają więc wykładniczo ze stałą czasową równą τ. Z tego względu stałą tę nazywamy *czasem relaksacji zderzeniowej*[1].

Ponieważ, jak to już wyżej zaznaczono, przy każdym zderzeniu elektronu z siecią krystaliczną następuje raptowna i całkowicie przypadkowa zmiana wartości i kierunku prędkości elektronu, zatem w zderzeniu tym zanika prawie całkowicie uporządkowana składowa jego ruchu. Biorąc to pod uwagę dochodzimy do wniosku, że czas relaksacji zderzeniowej jest w przybliżeniu równy średniemu czasowi między kolejnymi zderzeniami elektronów z siecią, o którym była mowa w p. 10.2.

Przejdźmy teraz do wyznaczenia prędkości unoszenia elektronów w stanie ustalonym. W tych warunkach $\frac{dv_{nu}}{dt} = 0$, a zatem

$$v_{nu} = -\frac{q\tau}{m_n^*} E \tag{10/27}$$

Jeżeli pole elektryczne jest słabe, to wywołuje ono niewielkie zaburzenie rozkładu prędkości termicznych elektronów. Można więc przyjąć, że średni czas między kolejnymi zderzeniami elektronu z siecią nie ulega zmianie. Pod tym założeniem prędkość unoszenia elektronów jest proporcjonalna do natężenia pola elektrycznego, a ruchliwość wyrażona wzorem

$$\mu_n = \frac{q\,\tau}{m_n^*} \tag{10/28}$$

jest wielkością niezależną od E. W ten sposób udowodniliśmy prawo Ohma, do którego uczony ten doszedł na podstawie przeprowadzonych doświadczeń.

Ważnym wnioskiem wypływającym ze wzoru (10/28) jest stwierdzenie, że ruchliwość jest proporcjonalna do czasu relaksacji τ. Czas ten zależy z kolei od stopnia niedoskonałości sieci krystalicznej, a więc od jej temperatury oraz koncentracji domieszek i innych defektów strukturalnych. Każdy rodzaj niedoskonałości charakteryzuje się przy tym nieco odmiennym mechanizmem rozpraszania elektronów i inną wartością czasu relaksacji.

Jeżeli w półprzewodniku występuje jednocześnie kilka przyczyn rozpraszania elektronów o czasach relaksacji τ_i ($i = 1, 2, \ldots$), to efektywny czas relaksacji występujący we wzorze (10/28) można obliczyć w następujący sposób: Prawdopodobieństwo tego, że nośnik ulegnie w czasie dt zderzeniu z centrum rozpraszania o czasie relaksacji τ_i wynosi dt/τ_i. Jeżeli poszczególne procesy rozpraszania są niezależne od siebie, to wypadkowe prawdopodobieństwo zde-

[1] Proces opisany wzorem (10/26) nazywamy relaksacją *zderzeniową* w odróżnieniu od innych procesów relaksacyjnych np. od omawianych w poprzednim rozdziale zjawisk rekombinacyjno-generacyjnych.

rzenia będzie sumą poszczególnych prawdopodobieństw, a zatem

$$\frac{dt}{\tau} = \sum_i \frac{dt}{\tau_i} \tag{10/29}$$

Stąd wynika następujący wzór na efektywny czas relaksacji zderzeniowej:

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \ldots \tag{10/30}$$

Na podstawie tego wzoru można stwierdzić, że decydujące znaczenie ma proces o najkrótszym czasie relaksacji, tzn. proces o największej intensywności rozpraszania.

Rozważania powyższe dotyczyły elektronów, jednakże wszystkie wyciągnięte z nich wnioski można rozciągnąć także na dziury.

W półprzewodnikach istnieje kilka mechanizmów rozpraszania nośników. W materiałach monokrystalicznych najważniejszą rolę odgrywa rozpraszanie wynikające z drgań termicznych sieci krystalicznej (jest to tzw. *rozpraszanie na fononach*, traktuje się je bowiem jako wynik zderzeń nośników z fononami) oraz *rozpraszanie na zjonizowanych domieszkach*. W tym ostatnim przypadku intensywność rozpraszania nośników, a więc również ich ruchliwość, zależą od sumy koncentracji zjonizowanych donorów i akceptorów, podczas gdy np. o koncentracji nośników decyduje, jak wynika z rozważań zawartych w rozdz. 8, różnica $N_d - N_a$. Tak więc ruchliwość nośników w półprzewodniku skompensowanym może być znacznie mniejsza niż ich ruchliwość w półprzewodniku samoistnym.

Domieszki nie zjonizowane powodują także zakłócenia periodyczności sieci krystalicznej. Wywołane przez nie rozpraszanie nośników jest jednakże słabe i jest zwykle do pominięcia wobec dwóch poprzednio wymienionych.

Ograniczając się do dwóch najważniejszych mechanizmów rozpraszania, można na podstawie wzoru (10/30) napisać następujący wzór na ruchliwość:

$$\mu = \left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right)^{-1} \tag{10/31}$$

gdzie μ_1 i μ_2 są to ruchliwości, które uzyskalibyśmy, gdyby poszczególne procesy rozpraszania nośników występowały samodzielnie.

Dla kryształów o wiązaniu kowalencyjnym (np. dla Ge i Si) można znaleźć w literaturze następujące wzory teoretyczne na ruchliwość μ_1 wynikającą z rozpraszania na fononach akustycznych oraz μ_2 wynikającą z rozpraszania na zjonizowanych domieszkach[1]:

$$\mu_1 = C_1(m^*)^{-5/2}T^{-3/2} \tag{10/32}$$

$$\mu_2 = \frac{C_2(m^*)^{1/2}}{N_d+N_a}T^{3/2} \tag{10/33}$$

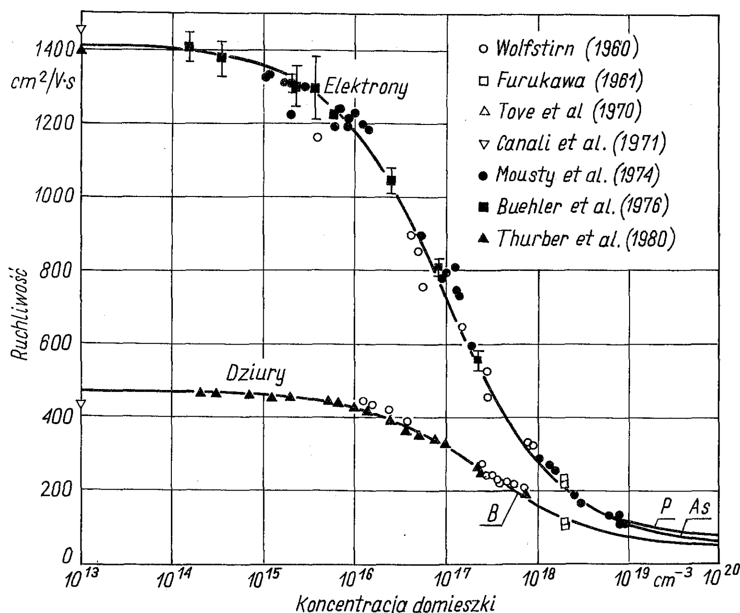
[1] Zob. np.: S. M. Sze: *Physics of semiconductor devices*. J. Wiley, New York 1981, § 1.5.

gdzie: C_1, C_2 – współczynniki zależne od rodzaju półprzewodnika; m^* – masa efektywna nośników danego rodzaju[1].

Przydatność tych wzorów jest niewielka, ponieważ występujące w nich parametry materiałowe nie są dostatecznie dobrze znane, a co ważniejsze – wzory te nie oddają we właściwy sposób zależności ruchliwości od temperatury.

Dla niektórych materiałów, a przede wszystkim dla krzemu, istnieją dość dokładne dane doświadczalne dotyczące ruchliwości elektronów i dziur przy różnych koncentracjach domieszek i w różnych temperaturach.

Przykładowe zestawienie takich danych opublikowanych w latach 1960–1980 przedstawiono na rys. 10/7. Na podstawie tego rysunku można stwierdzić, że przy małych koncentracjach domieszek (np. mniejszych niż 10^{15} cm^{-3}) ruchliwości są prawie niezależne od tych koncentracji. Świadczy to o tym, że w słabo domieszkowanym półprzewodniku decydujące znaczenie ma rozpraszanie nośników na fononach. W miarę powiększania koncentracji domieszek wzrasta ich udział w procesach rozpraszania i wskutek tego ruchliwości maleją.



10/7 Zależność ruchliwości elektronów i dziur w krzemie od koncentracji domieszek. Rodzaj domieszki: As — arsen, B — bor, P — fosfor
(Zestawienie danych doświadczalnych zaczerpnięto z: R. M. Warner, Jr., B. L. Grung: *Transistors, fundamentals for the integrated-circuit engineer.* J. Wiley, New York 1983. Krzywe obliczono na podstawie wzorów (10/34) — (10/36)

Problem 10.3. Na podstawie danych przedstawionych na rys. 10/7 wykreśl przebieg zależności ruchliwości μ_2 dla elektronów od koncentracji domieszek. Czy i w jakim zakresie zależność ta jest zgodna z przewidywaną wzorem (10/33). Przedyskutuj otrzymane wyniki.

---

[1] Jest to tzw. *masa efektywna przewodności*, która, podobnie jak masa efektywna gęstości stanów, jest pewną wielkością uśrednioną, różni się jednak od tej ostatniej sposobem uśredniania.

Analiza rysunku 10/7 pozwala stwierdzić, że dokładność z jaką znamy ruchliwości nośników nie jest duża. Wniosek ten dotyczy również innych parametrów materiałowych półprzewodników. Z tego względu stosowanie w obliczeniach, w których występują te parametry, większych dokładności niż do 2–3 cyfr znaczących jest niecelowe.

Na podstawie uzyskanych danych doświadczalnych zostały opracowane przez różnych autorów wzory empiryczne przedstawiające zależność ruchliwości elektronów i dziur od koncentracji domieszek (i ewentualnie temperatury). Dotyczą one przede wszystkim krzemu, który jest podstawowym materiałem wykorzystywanym w elementach półprzewodnikowych. Spośród tych wzorów najbardziej wiarygodne wydają się następujące zależności dotyczące krzemu w temperaturze 300 K[1]

– dla krzemu domieszkowanego fosforem

$$\mu_n = 68{,}5 + \frac{1345{,}5}{1+\left(\dfrac{9{,}20\cdot 10^{16}}{n}\right)^{-0{,}711}} - \frac{56{,}1}{1+\left(\dfrac{3{,}41\cdot 10^{20}}{n}\right)^{1{,}98}}\ [\mathrm{cm^2/V\cdot s}] \qquad (10/34)$$

– dla krzemu domieszkowanego arsenem

$$\mu_n = 52{,}2 + \frac{1364{,}8}{1+\left(\dfrac{9{,}68\cdot 10^{16}}{n}\right)^{-0{,}680}} - \frac{43{,}4}{1+\left(\dfrac{3{,}43\cdot 10^{20}}{n}\right)^{2}}\ [\mathrm{cm^2/V\cdot s}] \qquad (10/35)$$

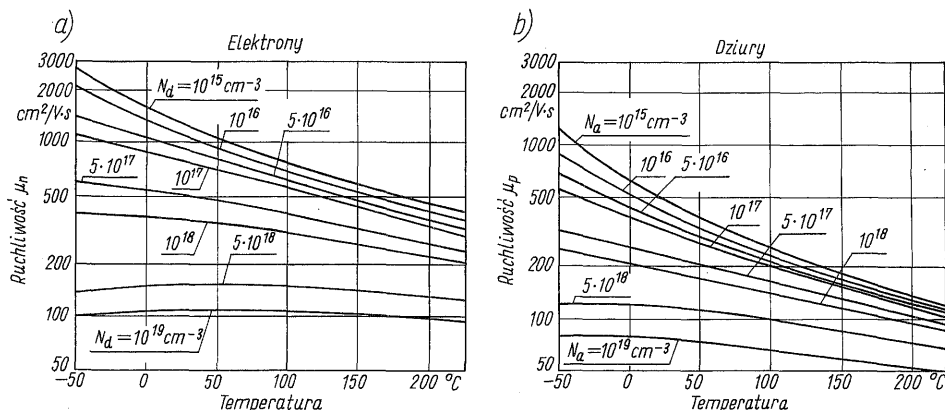
– dla krzemu domieszkowanego borem

$$\mu_p = 44{,}9\exp\left(-\frac{9{,}23\cdot 10^{16}}{p}\right) + \frac{470{,}5}{1+\left(\dfrac{2{,}23\cdot 10^{17}}{p}\right)^{-0{,}719}} - \frac{29{,}0}{1+\left(\dfrac{6{,}10\cdot 10^{20}}{p}\right)^{2}}\ [\mathrm{cm^2/V\cdot s}] \qquad (10/36)$$

Dwa pierwsze z tych wzorów są słuszne w zakresie koncentracji elektronów od 10^{13} do $5\cdot 10^{21}$ cm^{-3}, trzeci zaś w zakresie koncentracji dziur do $1{,}2\cdot 10^{21}$ cm^{-3}.

Linie ciągłe na rys. 10/7 przedstawiają zależności obliczone na podstawie podanych wyżej wzorów empirycznych przy założeniu, że koncentracje nośników większościowych są równe koncentracjom odpowiednich domieszek. Przy

[1] Według G. Masetti, M. Severi, S. Solmi: Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon. *IEEE Trans. on Electron Devices*, **Ed-20**, 764 (1983).



10/8 Zależność ruchliwości elektronów (a) i dziur (b) w krzemie od temperatury przy różnych koncentracjach domieszek
(Zaczerpnięto z: D. J. Hamilton, W. G. Howard: *Basic integrated circuit engineering*. McGraw--Hiii, New York 1975)

koncentracjach domieszek nie przekraczających wartości 10^{19} cm^{-3} trzeci składnik we wzorach (10/34)–(10/36) może być zaniedbany jako pomijalnie mały wobec dwóch pozostałych. Również w tym zakresie koncentracji wpływ rodzaju domieszki donorowej jest prawie niedostrzegalny.

Wpływ temperatury na ruchliwość nośników w krzemie przy różnych koncentracjach domieszek przedstawiono na rys. 10/8.

W materiałach słabo domieszkowanych o ruchliwości decyduje, jak to już wyżej zaznaczono, rozpraszanie nośników na fononach. Ze wzrostem temperatury zwiększa się energia drgań sieci krystalicznej (czyli powiększa się koncentracja fononów), a zatem ruchliwość maleje. Zależność ta nie przebiega jednak dokładnie wg wzoru (10/32). W szerokim zakresie temperatur można przedstawić ją za pomocą wzoru empirycznego

$$\mu(T) = \mu(300)\left(\frac{T}{300}\right)^{-m} \tag{10/37}$$

gdzie: $\mu(T)$ – ruchliwość danego rodzaju nośników w temperaturze T; m – wykładnik zależny od rodzaju półprzewodnika i rodzaju nośników (zob. tabl. 10/2).

Korzystając z danych zawartych w tabl. 10/1 i 10/2, otrzymuje się dla słabo domieszkowanego krzemu następujące wzory liczbowe słuszne w zakresie temperatur 200–500 K (tzn. od -70 ... do $+230$°C):

$$\mu_n = 1{,}40\cdot 10^9\, T^{-2{,}42}\ [\mathrm{cm^2/V\cdot s}] \tag{10/38}$$

$$\mu_p = 1{,}32\cdot 10^8\, T^{-2{,}2}\ [\mathrm{cm^2/V\cdot s}] \tag{10/39}$$

Przy większych koncentracjach domieszek obok rozpraszania nośników na fononach istotną rolę odgrywa rozpraszanie na zjonizowanych domieszkach. Efektywność tego rozpraszania jest tym większa, im mniejsze są prędkości termiczne nośników, tzn. im niższa temperatura. Ze wzrostem temperatury ma-

10/2 **Wartości wykładnika *m* ze wzoru (10/37) dla słabo domieszkowanych półprzewodników w zakresie temperatur 200-500 K**

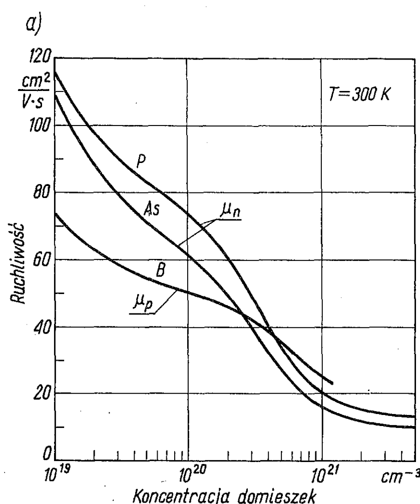
Półprzewodnik	Rodzaj nośników		Źródło
	Elektrony	Dziury	
Si	2,42	2,20	1)
Ge	1,66	2,33	1)
GaAs	1,0	2,1	1)
GaSb	2,0	0,9	2)
InSb	1,6	2,1	2)

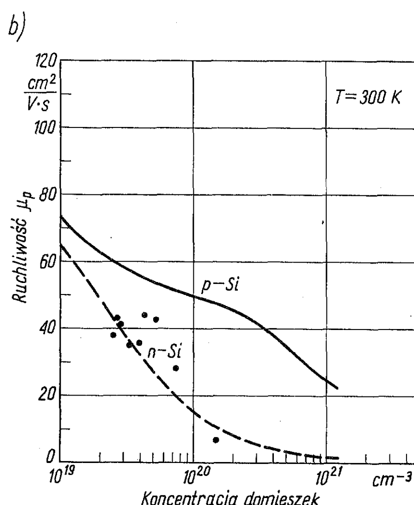
1) S. M. Sze: *Physics of semiconductor devices*. J. Wiley, New York 1981, s. 28–29.
2) J. M. Feldman: *The physics and circuit properties of transistors*. J. Wiley, New York 1972, s. 194.

leje więc μ_1, zwiększa się natomiast μ_2 – zob. wzór (10/33). W wyniku tego w miarę powiększania koncentracji domieszek obserwuje się coraz słabszy wpływ temperatury na ruchliwość nośników – zob. rys. 10/8.

Przy bardzo dużych koncentracjach domieszek (np. większych niż 10^{19} cm^{-3}) uwidoczniają się następujące efekty, które przy mniejszych koncentracjach są praktycznie niedostrzegalne:

a) obok rozpraszania nośników na zjonizowanych domieszkach istotną rolę zaczyna odgrywać rozpraszanie wywołane zderzeniami z innymi nośnikami – jest to tzw. *rozpraszanie na swobodnych nośnikach*. Wywołuje ono raptowny spadek ruchliwości po przekroczeniu wartości koncentracji domieszek rzędu $2 \cdot 10^{20}$ cm^{-3} – zob. rys. 10/9a. Ten dodatkowy





10/9 Ruchliwość nośników w bardzo silnie domieszkowanym krzemie: a) nośniki większościowe (rodzaj domieszki: As — arsen, B — bor, P — fosfor) — obliczono na podstawie wzorów (10/34)—(10/36); b) dziury w krzemie typu *p* (linia ciągła) i typu *n* (linia przerywana oraz punkty doświadczalne) — opracowano na podstawie: W. Kuźmicz: Zjawiska występujące w półprzewodnikach silnie domieszkowanych i ich wpływ na parametry przyrządów półprzewodnikowych. Zesz. *Naukowe Politechniki Warszawskiej*, Elektronika. Z. 61 (1985)

mechanizm rozpraszania nośników jest reprezentowany przez ostatni składnik we wzorach (10/34) — (10/36);

b) ruchliwość nośników w półprzewodniku danego typu staje się zależna do rodzaju domieszki, np. w krzemie domieszkowanym fosforem jest ona większa niż w krzemie zawierającym domieszkę arsenową;

c) przy koncentracjach domieszek większych niż $3 \cdot 10^{20}$ cm^{-3} ruchliwość dziur w krzemie typu *p* jest większa niż ruchliwość elektronów w krzemie typu *n*;

d) ruchliwość nośników danego rodzaju zależy od tego czy są one nośnikami większościowymi, czy też mniejszościowymi. Pokazano to na rys. 10/9b, na którym przedstawiono ruchliwość dziur w krzemie typu *p* i typu *n*. Ruchliwość dziur „większościowych" wyliczono ze wzoru (10/36), natomiast ruchliwości dziur „mniejszościowych" zostały zestawione na podstawie pomiarów wykonanych przez różnych autorów. Do niedawna przyjmowano powszechnie, że ruchliwość nośników mniejszościowych nie różni się od ruchliwości nośników większościowych tego samego rodzaju przy tej samej koncentracji domieszek. Stwierdzenie to znajdowało potwierdzenie doświadczalne przy koncentracjach domieszek mniejszych niż 10^{19} cm^{-3}. Ostatnio zauważono, że w bardzo silnie domieszkowanym krzemie ruchliwość dziur „mniejszościowych" jest kilkakrotnie mniejsza niż ruchliwość dziur „większościowych". Dla elektronów brak jest dotychczas danych dotyczących ich ruchliwości w bardzo silnie domieszkowanym krzemie typu *p*.

Szczególnie silne rozpraszanie nośników występuje w półprzewodnikach polikrystalicznych stosowanych w niektórych elementach półprzewodnikowych. Granice ziaren krystalicznych oraz liczne defekty strukturalne zmniejszają ruchliwość nośników do wartości porównywalnych z ruchliwościami spotykanymi w analogicznych materiałach monokrystalicznych przy bardzo dużych koncentracjach domieszek.

Także ruchliwość elektronów w metalach jest znacznie mniejsza niż w słabo domieszkowanych półprzewodnikach monokrystalicznych (zob. problem 10.2). Maleje ona natomiast, podobnie jak w półprzewodnikach, ze wzrostem temperatury. Świadczy to o tym, że w metalach największe znaczenie ma rozpraszanie elektronów na fononach. Dopiero przy temperaturach mniejszych niż 20 K istotną rolę zaczyna odgrywać rozpraszanie na defektach strukturalnych. Uwidocznia się ono na przebiegu zależności oporu właściwego od temperatury – zob. rys. 10/12.

10.5 PRZEWODNOŚĆ WŁAŚCIWA

Przewodność właściwa półprzewodnika zależy od koncentracji występujących w nim nośników i od ich ruchliwości – zob. wzór (10/23). Dla półprzewodnika samoistnego wzór ten upraszcza się do postaci

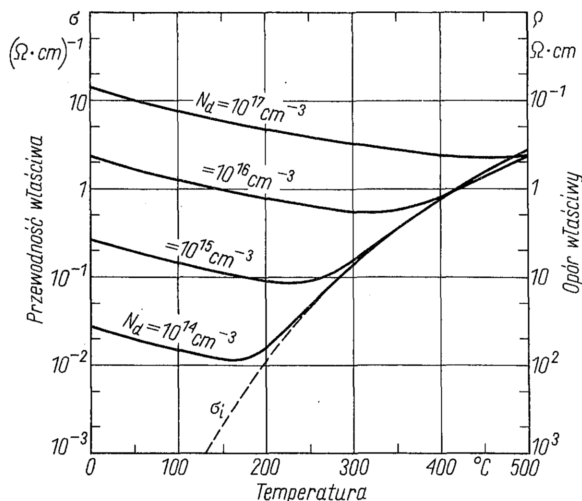
$$\sigma_i = qn_i(\mu_n + \mu_p) \tag{10/40}$$

Ze względu na to, że koncentracja samoistna n_i jest bardzo szybko rosnącą funkcją temperatury, przewodność samoistna σ_i wzrasta również szybko z temperaturą – zob. rys. 10/10.

Dla półprzewodnika typu *n* (spełniającego warunki $N_d \gg N_a$ oraz $N_d \gg n_i$) można we wzorze (10/23) położyć $n = N_d$ oraz $p = 0$. Otrzymujemy wówczas

$$\boxed{\sigma = qN_d\mu_n} \tag{10/41}$$

Jedyną wielkością zależną od temperatury jest tu ruchliwość elektronów μ_n. Na podstawie rozważań zawartych w poprzednim punkcie można stwierdzić,



10/10 Zależność przewodności właściwej krzemu typu *n* od temperatury przy różnych koncentracjach donorów

że w interesującym nas zakresie temperatur przewodność półprzewodników domieszkowanych nieznacznie maleje ze wzrostem temperatury.

Na rysunku 10/10 przedstawiono zależność przewodności właściwej (i oporu właściwego) krzemu od temperatury przy różnych wartościach koncentracji donorów. Linia przerywana przedstawia przewodność samoistną wyznaczoną ze wzoru (10/40), linie ciągłe zaś – przewodności obliczone na podstawie wzoru (10/41) w zakresie słuszności tego wzoru, a z dokładniejszych wzorów – poza tym zakresem.

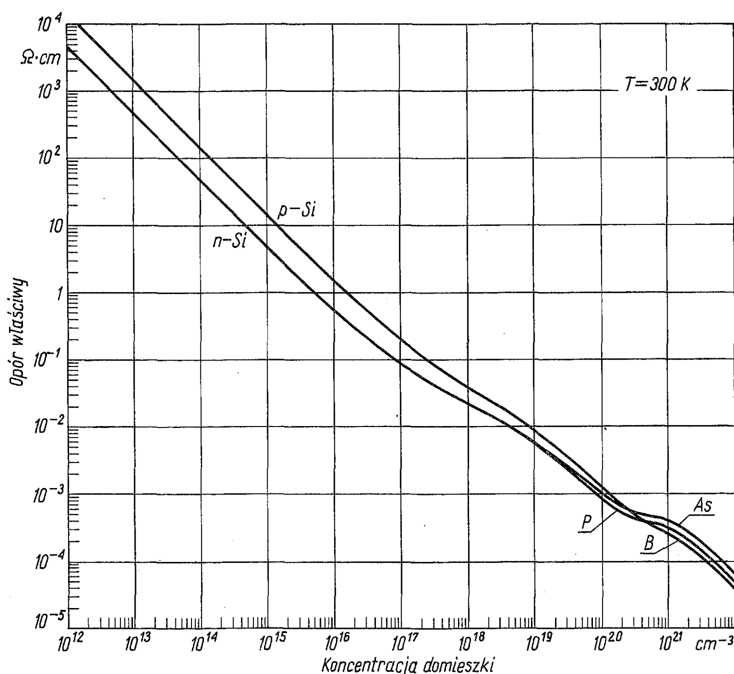
Dla półprzewodnika typu *p* analogiczny wzór na przewodność ma postać:

$$\boxed{\sigma = qN_a \mu_p} \qquad (10/42)$$

Problem 10.4. Wyznacz przewodność właściwą, opór właściwy oraz współczynnik temperaturowy przewodności dla krzemu domieszkowanego borem o koncentracji $5 \cdot 10^{16}$ cm^{-3} w temperaturze pokojowej. Należy przyjąć, że ruchliwość nośników zmienia się z temperaturą zgodnie ze wzorem (10/37); wyznacz wyładnik *m* na podstawie wykresów podanych na rys. 10/8.

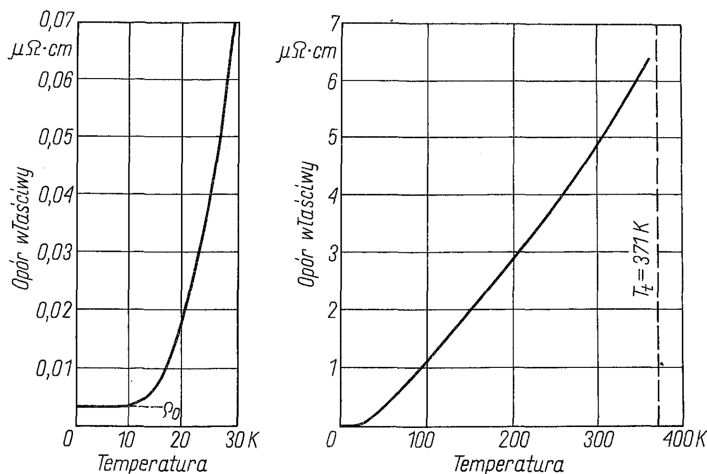
Na rysunku 10/11 przedstawiono zależność oporu właściwego krzemu od koncentracji domieszek w temperaturze pokojowej. Obliczono go przy założeniu, że koncentracje nośników większościowych są równe koncentracjom domieszek. Ruchliwości nośników wyznaczono ze wzorów 10/34–10/36.

Koncentracje domieszek w elementach półprzewodnikowych mogą zmieniać się w bardzo szerokich granicach, dochodzących nawet do 8 dekad. Różne obszary elementu półprzewodnikowego mogą więc wykazywać opory właściwe, które w skrajnych przypadkach mają się do siebie jak jeden do kilku milionów.



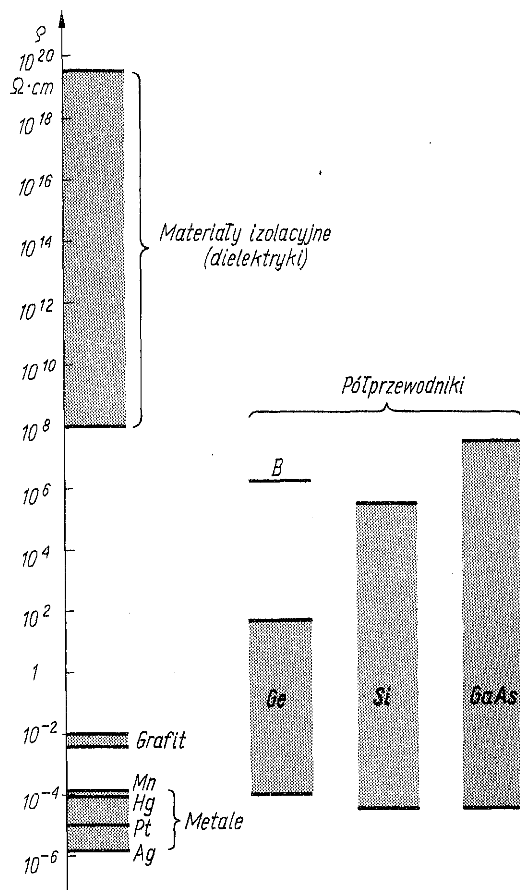
10/11 Zależność oporu właściwego krzemu od koncentracji domieszek. Rodzaj domieszki: As — arsen, B — bor, P — fosfor

W metalach występuje tylko jeden rodzaj nośników ładunku – elektrony, przy czym ich koncentracja nie zależy od temperatury. Opór właściwy metali zmienia się zatem z temperaturą wskutek tego, że ruchliwość elektronów jest funkcją temperatury. Przykład zależności oporu właściwego metalu od temperatury pokazano na rys. 10/12.



10/12 Zależność oporu właściwego sodu od temperatury ϱ_0 — tzw. *opór szczątkowy*, wynikający z rozpraszania elektronów na defektach strukturalnych; T_t — temperatura topnienia

Orientacyjne wartości oporu właściwego metali, materiałów izolacyjnych oraz kilku półprzewodników zestawiono na rys. 10/13. Z wykresu tego wynika, że w najsilniej domieszkowanych półprzewodnikach uzyskuje się wartości oporu właściwego takie, jak w gorzej przewodzących metalach. Z kolei półprzewodniki skompensowane mogą pod pewnymi względami konkurować z materiałami izolacyjnymi[1].



10/13 Opory właściwe metali, półprzewodników i materiałów izolacyjnych

10.6 ZJAWISKA WYSTĘPUJĄCE PRZY DUŻYCH NATĘŻENIACH POLA ELEKTRYCZNEGO

10.6.1* Zależność ruchliwości od natężenia pola elektrycznego

Rozważania zawarte w poprzednim punkcie dotyczyły zakresu słabych pól elektrycznych. W tych warunkach można było przyjmować, że zaburzenie rozkładu prędkości termicznych nośników wywołane działaniem pola elektrycz-

[1] Przykładem może tu być *półizolacyjny* arsenek galu, o którym była mowa w p. 7.6.3.

nego jest niewielkie, w wyniku czego ruchliwość nośników jest niezależna od natężenia tego pola. Przykładowo, przy $E = 500$ V/cm prędkość unoszenia elektronów w słabo domieszkowanym krzemie jest rzędu $5 \cdot 10^5$ cm/s, co stanowi ok. 5% ich prędkości termicznej (która, jak pamiętamy, jest rzędu 10^7 cm/s). Przy większych wartościach natężenia pola elektrycznego prędkości unoszenia nośników stają się porównywalne z ich prędkościami termicznymi, a zatem ruchy termiczne i ruchy unoszenia nie mogą być traktowane jako niezależne od siebie.

W miarę wzrostu natężenia pola elektrycznego powiększa się średnia energia nośników. Można więc uważać, że gaz elektronowy ma wyższą temperaturę niż sieć krystaliczna. Takie elektrony nazywamy elektronami *gorącymi*, a przypisaną im temperaturę nazywamy *temperaturą elektronową.*

Z warunku równości energii pobieranej przez elektrony z pola elektrycznego i energii przekazywanej sieci, otrzymuje się dla półprzewodników o wiązaniu kowalencyjnym następujący wzór na temperaturę elektronową:

$$\frac{T_e}{T} = \frac{1}{2}\left(1 + \sqrt{1 + \frac{3\pi}{8}\left(\frac{\mu_0 E}{C_s}\right)^2}\right) \tag{10/43}$$

gdzie: T – temperatura sieci krystalicznej; μ_0 – ruchliwość nośników przy małych wartościach natężenia pola E; C_s – prędkość dźwięku w rozpatrywanym krysztale.

Im wyższa jest temperatura elektronowa nośników, tym bardziej efektywnie tracą one energię na rzecz sieci (powiększa się bowiem częstość zderzeń i wartość energii przekazywanej w każdym zderzeniu). W wyniku tego prędkość unoszenia przestaje być proporcjonalna do natężenia pola, a przy dalszym powiększaniu E osiąga ona pewną wartość graniczną, nazywaną *prędkością nasycenia* – zob. rys. 10/3.

Prędkość ta zależy od rodzaju nośników i rodzaju półprzewodnika (zob. tabl. 10/1) oraz temperatury – maleje mianowicie z jej wzrostem, prawie nie zależy natomiast od koncentracji domieszek.

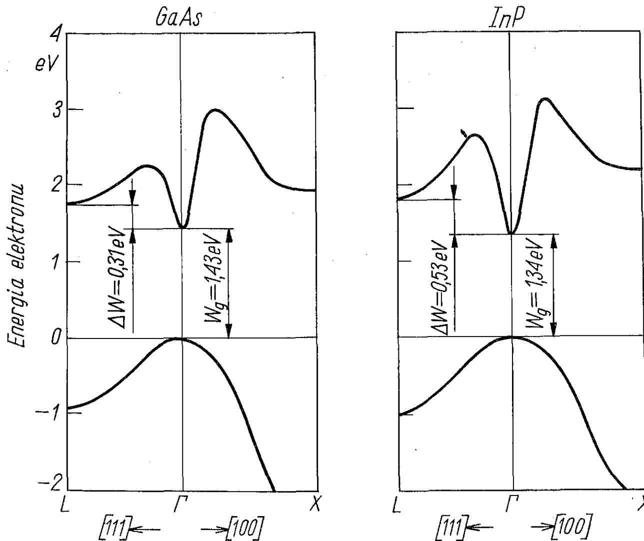
Zależność ruchliwości nośników w takich półprzewodnikach jak np. krzem lub german od natężenia pola elektrycznego można wyrazić za pomocą następującego wzoru empirycznego:

$$\mu = \mu_0 \left[\frac{1}{1 + \left(\frac{\mu_0 E}{v_s}\right)^\beta}\right]^{1/\beta} \tag{10/44}$$

gdzie: μ_0 – ruchliwość rozpatrywanego rodzaju nośników przy małych wartościach natężenia pola E; v_s – ich prędkości nasycenia; β – wykładnik, którego wartość wynosi np. dla krzemu 1,4–1,5.

10.6.2* Efekt Gunna

W niektórych półprzewodnikach, np. w GaAs i InP, przebieg zależności prędkości unoszenia elektronów od natężenia pola elektrycznego jest bardziej skomplikowany. Na modelu pasmowym obu tych półprzewodników (rys. 10/14) można zauważyć, że pasmo przewodnictwa zawiera dolinę centralną w środku strefy Brillouina oraz tzw. *doliny boczne* położone na osiach [111]. Dolina centralna charakteryzuje się małą masą efektywną i w związku z tym dużą ruchliwością elektronów (np. dla GaAs: $m_n^* = 0{,}067\, m_0$; $\mu_n = 8500\ \mathrm{cm^2/V\cdot s}$), natomiast doliny boczne – dużą masą efektywną i małą ruchliwością elektronów (dla GaAs: $m_n^* = 0{,}55\, m_0$, $\mu_n \approx 100\ \mathrm{cm^2/V\cdot s}$).



10/14 Struktura energetyczna arsenku galu i fosforku indu (T = 300 K)

W skali energetycznej doliny boczne są przesunięte o ΔW w stosunku do doliny centralnej, czyli w stosunku do dna pasma przewodnictwa. Przesunięcie to wynosi np. dla GaAs 0,31 eV. W stanie równowagi termodynamicznej elektrony obsadzają głównie stany należące do doliny centralnej, pozostawiając doliny boczne praktycznie nie obsadzone. Przy małych wartościach natężenia pola elektrycznego elektrony ulegają rozpraszaniu, pozostając w obrębie doliny centralnej (jest to tzw. *rozpraszanie wewnątrzdolinowe*, w którym biorą udział fonony akustyczne o niewielkich energiach), natomiast w miarę powiększania natężenia pola wzrasta prawdopodobieństwo przejść elektronów z doliny centralnej do dolin bocznych (tzw. *rozpraszanie międzydolinowe*, w którym biorą udział głównie fonony optyczne o stosunkowo dużych energiach).

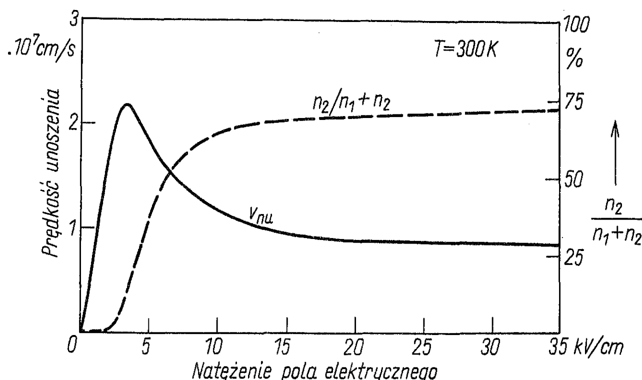
W wyniku działania sił pola elektrycznego maleje więc obsadzenie stanów o niższych energiach w dolinie centralnej, rośnie natomiast obsadzenie stanów o wyższych energiach w dolinach bocznych. Ponieważ te ostatnie charaktery-

zują się mniejszą ruchliwością elektronów, zatem, po przekroczeniu pewnej krytycznej wartości natężenia pola E_{kr}, średnia prędkość unoszenia elektronów zaczyna się zmniejszać. Prędkość ta jest średnią ważoną prędkości wynikających z ruchliwości elektronów w dolinie centralnej i w dolinach bocznych, wynosi więc

$$v_{nu} = \left(\frac{n_1 \mu_1 + n_2 \mu_2}{n_1 + n_2}\right) E \tag{10/45}$$

gdzie: n_1, n_2 oznaczają odpowiednio koncentracje elektronów w dolinie centralnej i w dolinach bocznych; μ_1, μ_2 zaś – ruchliwości tych elektronów (zależne od natężenia pola elektrycznego).

Na rysunku 10/15 pokazano, jak ze wzrostem natężenia pola zmienia się obsadzenie stanów w dolinach bocznych oraz prędkość unoszenia elektronów. W pewnym zakresie natężeń pola charakterystyka v_{nu} (E) ma przebieg opadający, tzn. dodatnim przyrostom E odpowiadają ujemne przyrosty v_{nu}.



10/15 **Zależność prędkości unoszenia elektronów od natężenia pola elektrycznego w arsenku galu (linia ciągła). Linia przerywana przedstawia stosunek n_2/n_1+n_2 wyrażony w procentach**

O zachowaniu się danego elementu w obwodzie prądu zmiennego decydują nie tylko jego parametry statyczne (takie np. jak ruchliwość zdefiniowana wzorem (10/12)), lecz również parametry *przyrostowe* (różniczkowe), w definicji których występują małe przyrosty wielkości zmieniających się w czasie. Wprowadzając pojęcie ruchliwości różniczkowej

$$\mu_r = \frac{dv_{nu}}{dE} \tag{10/46}$$

można stwierdzić na podstawie rys. 10/15, że przy natężeniach pola większych od E_{kr} jest ona ujemna. Dla GaAs $E_{kr} = 3{,}2$ kV/cm, a maksymalna ujemna wartość ruchliwości różniczkowej wynosi -2400 cm^2/V·s. Ujemna ruchliwość różniczkowa pociąga za sobą ujemną wartość przewodności różniczkowej i ujemny opór różniczkowy rozpatrywanego obszaru półprzewodnika.

Dominującą cechą oporu dodatniego jest, jak wiadomo, rozpraszanie energii prądu elektrycznego. Opór taki dołączony do źródła sygnału elektrycznego pobiera energię z tego źródła, czyli, jak się często mówi, *obciąża* je.

W przypadku oporu ujemnego przepływ energii prądu zmiennego zachodzi w przeciwnym kierunku. Opór ten „*odciąża*" więc źródło, czyli zmniejsza straty energii w obwodzie. Przy spełnieniu określonych warunków może wówczas nastąpić wzbudzenie się nietłumionych drgań elektrycznych.

Elementy półprzewodnikowe, których działanie opiera się na wykorzystaniu opadającego odcinka charakterystyki $v_{nu}(E)$ do generacji drgań elektrycznych, nazywamy *diodami Gunna*[1], a samo zjawisko wzbudzania się tych drgań nosi nazwę *efektu Gunna*.

10.6.3* Jonizacja zderzeniowa

W punkcie 10.3, przy omawianiu zachowania się nośników ładunku w sieci krystalicznej, stwierdziliśmy, że w kolejnych zderzeniach nośniki przekazują sieci określone porcje energii (zob. rys. 10/4). Przy dostatecznie dużym natężeniu pola elektrycznego nośnik może między kolejnymi zderzeniami uzyskać energię wystarczającą do zjonizowania sieci, tzn. do wytworzenia pary elektron-dziura. Zjawisko to, nazywane *jonizacją zderzeniową*, jest procesem odwrotnym w stosunku do rekombinacji Augera, o której była mowa w p. 9.5.

W procesie Augera energia wyzwalająca się przy rekombinacji pary elektron-dziura zostaje przekazana nośnikowi ładunku (elektronowi lub dziurze), dzięki czemu powiększa się jego energia kinetyczna. Przy jonizacji zderzeniowej natomiast energia kinetyczna elektronu lub dziury zostaje zużyta na generację pary elektron-dziura. Energia ta przekazana w pojedynczym zderzeniu z siecią krystaliczną musi być oczywiście większa niż przerwa energetyczna danego półprzewodnika – zob. rys. 10/16.

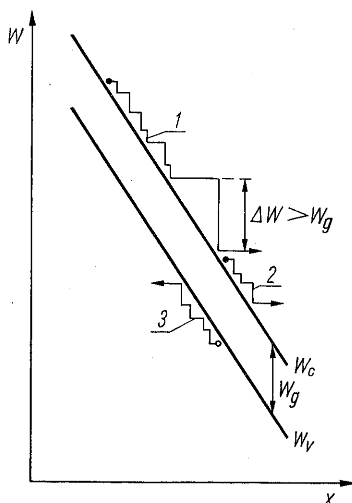
Zjawisko jonizacji zderzeniowej można opisać, od strony ilościowej, za pomocą szybkości generacji par elektron-dziura, którą można wyrazić wzorem

$$g = \alpha_n n v_{nu} + \alpha_p p v_{pu} \tag{10/47}$$

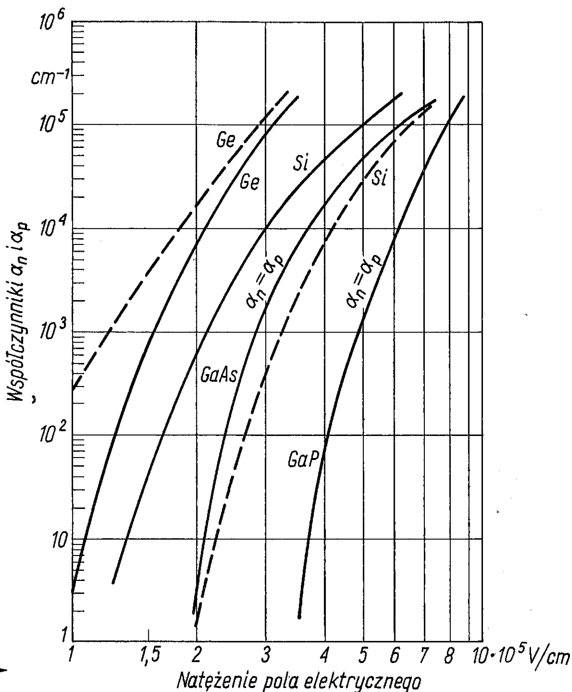
gdzie: α_n i α_p są to tzw. *współczynniki jonizacji zderzeniowej* równe średniej liczbie par elektron-dziura wytwarzanych przez dany rodzaj nośnika na drodze równej jednostce długości przebytej w kierunku pola elektrycznego. Wielkości te zależą bardzo silnie od natężenia pola elektrycznego, a ponadto od rodzaju nośników, rodzaju półprzewodnika i temperatury.

Przebieg zależności współczynników α_n i α_p od natężenia pola dla kilku półprzewodników pokazano na rys. 10/17. Można tam zauważyć, że na ogół

[1] W języku angielskim elementy te są określane jako „transferred-electron devices", co dosłownie oznacza elementy z przeniesionymi elektronami. Elementy te odgrywają bardzo ważną rolę w technice mikrofalowej. Znaczenie efektu Gunna wykracza jednakże poza te zastosowania.



10/16 Zjawisko jonizacji zderzeniowej w półprzewodniku
1 — elektron pierwotny;
2 — dodatkowy elektron wytworzony w wyniku jonizacji sieci krystalicznej; *3* — dziura wytworzona w wyniku jonizacji



10/17 Zależność współczynników jonizacji α_n (linie ciągłe) i α_p (linie przerywane) od natężenia pola elektrycznego dla różnych półprzewodników (Zaczerpnięto z: S. M. Sze: *Physics of semiconductor devices*. J. Wiley, New York 1969)

im większa jest przerwa energetyczna półprzewodnika, tym większe musi być natężenie pola, przy którym osiąga się daną wartość współczynnika jonizacji zderzeniowej.

10/3 Wartości współczynników *A*, *b*, *m* ze wzoru (10/45)

Półprzewodnik	Rodzaj nośników	*A* cm⁻¹	*b* V/cm	*m* —
Si	elektrony	$3{,}8 \cdot 10^6$	$1{,}75 \cdot 10^6$	1
	dziury	$2{,}25 \cdot 10^7$	$3{,}26 \cdot 10^6$	1
Ge	elektrony	$1{,}55 \cdot 10^7$	$1{,}56 \cdot 10^6$	1
	dziury	$1{,}0 \cdot 10^6$	$1{,}28 \cdot 10^6$	1
GaAs	elektrony	$1{,}34 \cdot 10^6$	$2{,}03 \cdot 10^6$	2
	dziury	$4{,}0 \cdot 10^5$	$1{,}8 \cdot 10^6$	2

Zależność współczynników α od natężenia pola można wyrazić następującym wzorem empirycznym:

$$\alpha = A \exp\left[-\left(\frac{b}{E}\right)^m\right] \tag{10/48}$$

gdzie: A, b, m – współczynniki zależne od rodzaju nośników, rodzaju półprzewodnika i temperatury. Wartości tych współczynników dla kilku półprzewodników w temperaturze pokojowej zestawiono w tabl. 10/3.

Ze wzrostem temperatury maleje średnia droga swobodna nośników, a wraz z nią średnia energia nabywana przez nośnik między kolejnymi zderzeniami z siecią. W wyniku tego maleją również współczynniki jonizacji.

10.7 PRĄD DYFUZYJNY

Obok prądów unoszenia występują w półprzewodniku prądy dyfuzyjne wywołane nierównomiernym rozkładem koncentracji lub energii nośników.

Dyfuzja jest to proces będący wyrazem dążenia natury do wyrównania rozkładu koncentracji i energii cząstek (atomów, molekuł, elektronów) w danym obszarze. Dobrze znanym przykładem dyfuzji jest rozpraszanie się dymu w powietrzu.

Procesy dyfuzji zachodzą wskutek ruchów cieplnych cząstek. W wyniku zachodzących przy tym zderzeń następują przypadkowe zmiany kierunku ruchu cząstek, wskutek czego przemieszczają się one stosunkowo powoli z obszarów o większej koncentracji (lub energii) do obszarów o mniejszej koncentracji (energii).

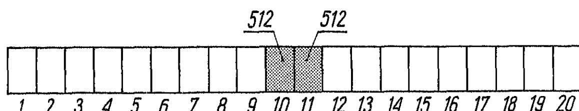
Dyfuzja nośników (a także innych cząstek) charakteryzuje się następującymi najbardziej istotnymi cechami:

1) odbywa się ona bez udziału jakichkolwiek sił zewnętrznych; jeżeli siły takie występują (np. wskutek istnienia pola elektrycznego), to wywołują one ruchy nośników całkowicie niezależne od procesu dyfuzji;

2) nośniki nie oddziałują na siebie, a więc np. odpływ nośników z obszarów o większej koncentracji nie następuje wskutek elektrostatycznego odpychania jednych nośników przez inne, lecz w wyniku ich całkowicie przypadkowych ruchów;

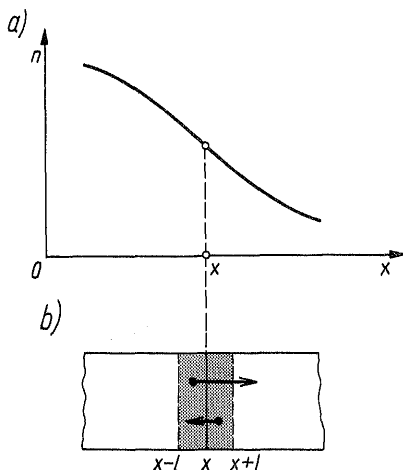
3) nośniki muszą być *wzbudzone* wskutek zderzeń z ośrodkiem, w którym się znajdują, i to w taki sposób, aby każdy z nich miał jednakowe prawdopodobieństwo poruszania się w którymkolwiek kierunku.

Problem 10.5. W celu lepszego zrozumienia mechanizmu dyfuzji rozpatrzmy następujący prosty model tego procesu w przestrzeni jednowymiarowej: przestrzeń składa się z 20 liniowo ułożonych komórek (rys. 10/18), przy czym w 2 środkowych komórkach znajduje się po 512 kulek. Załóżmy, że co sekundę połowa kulek z każdej komórki przesuwa się do sąsiedniej lewej komórki, druga połowa zaś – do sąsiedniej prawej komórki. Sporządź wykresy przedstawiające rozkłady kulek po upływie 1, 2, ..., 9 sekund.



10/18

Zajmiemy się teraz wyznaczeniem gęstości prądu dyfuzyjnego. W tym celu posłużymy się prostym modelem jednowymiarowym. Założymy mianowicie, że rozkład koncentracji elektronów w półprzewodniku jest opisany funkcją $n(x)$ – rys. 10/19a oraz że wszelkie ruchy elektronów (także ruchy termiczne) odbywają się w kierunku równoległym do kierunku osi x. Założymy również, że $T = \text{const}$, a zatem średnia energia ruchów cieplnych jest niezależna od x.



10/19
Przepływ nośników ładunku wywołany nierównomiernym rozkładem ich koncentracji: a) rozkład koncentracji elektronów w modelu jednowymiarowym; b) strumienie elektronów przechodzących przez płaszczyznę x

W celu znalezienia gęstości prądu elektronowego rozpatrzmy prostopadłościenną próbkę półprzewodnika o przekroju jednostkowym – rys. 10/19b. Liczba elektronów przechodzących przez płaszczyznę x w kierunku zgodnym z kierunkiem osi x w czasie τ równym średniemu czasowi między kolejnymi zderzeniami elektronu z siecią krystaliczną wynosi $\frac{1}{2}\, n(x-l)\, l$, gdzie l – średnia droga swobodna elektronu, a $n(x-l)$ – koncentracja elektronów w płaszczyźnie $x-l$. Czynnik 1/2 w tym wyrażeniu jest wynikiem uwzględnienia faktu, że podobnie jak w problemie 10.5, połowa elektronów znajdujących się po lewej stronie płaszczyzny x porusza się w prawo, druga połowa zaś – w lewo.

Analogiczna liczba elektronów przechodzących przez rozpatrywaną płaszczyznę w przeciwnym kierunku wynosi $\frac{1}{2}\, n(x+l)\, l$. Gęstość elektronowego prądu dyfuzyjnego jest proporcjonalna do różnicy tych liczb. Uwzględniając to, że elektron ma ładunek równy $-q$, znajdziemy, że szukana gęstość prądu wynosi

$$J_{nd}(x) = \frac{ql}{2\tau}\,[n(x+l)-n(x-l)] \qquad (10/49)$$

Przyjmując, że w rozważanym obszarze półprzewodnika rozkład koncentracji elektronów może być aproksymowany funkcją liniową współrzędnej x, możemy napisać

$$n(x+l) = n(x) + \left.\frac{\mathrm{d}n}{\mathrm{d}x}\right|_x l \qquad (10/50)$$

$$n(x-l) = n(x) - \left.\frac{\mathrm{d}n}{\mathrm{d}x}\right|_x l \qquad (10/51)$$

co po podstawieniu do (10/49) daje

$$J_{nd}(x) = \frac{ql^2}{\tau}\left.\frac{\mathrm{d}n}{\mathrm{d}x}\right|_x \qquad (10/52)$$

Średnia droga swobodna elektronu jest związana z jego prędkością termiczną v_{th} zależnością

$$l = v_{th}\,\tau \qquad (10/53)$$

Uwzględniając to, otrzymujemy, dla rozpatrywanego przypadku jednowymiarowego

$$J_{nd}(x) = q\, v_{th}^2\, \tau \left.\frac{\mathrm{d}n}{\mathrm{d}x}\right|_x \qquad (10/54)$$

Gęstość elektronowego prądu dyfuzyjnego jest więc proporcjonalna do *gradientu koncentracji* elektronów[1], przy czym współczynnik proporcjonalności jest zależny od intensywności ruchów termicznych elektronów w półprzewodniku.

Wzór (10/54) zapisuje się zwykle w postaci[2]

$$\boxed{J_{nd} = qD_n \frac{\mathrm{d}n}{\mathrm{d}x}} \qquad (10/55)$$

przy czym D_n nosi nazwę *współczynnika dyfuzji* elektronów. W rozpatrywanym tu modelu jednowymiarowym współczynnik ten równa się $v_{th}^2\tau$.

Jeżeli przy założeniu, że koncentracja elektronów zależy tylko od współrzędnej x, uwzględnimy to, iż ruchy termiczne tych nośników odbywają się

---

[1] *Gradient* jest to operator różniczkowy, oznaczany symbolem grad lub ∇, który działając na funkcję skalarną $\varphi(x, y, z)$ daje w wyniku wektor

$$\nabla\varphi = \frac{\partial\varphi}{\partial x}\boldsymbol{i} + \frac{\partial\varphi}{\partial y}\boldsymbol{j} + \frac{\partial\varphi}{\partial z}\boldsymbol{k}$$

gdzie: $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ – wektory jednostkowe prostokątnego układu współrzędnych x, y, z. Wektor ten określa kierunek i szybkość największego wzrostu funkcji w danym punkcie. W przypadku jednowymiarowym, gdy funkcja φ zależy tylko od jednej współrzędnej przestrzennej x

$$|\nabla\varphi| = \frac{\mathrm{d}\varphi}{\mathrm{d}x}$$

[2] Wzór ten jest szczególnym przypadkiem tzw. *pierwszego prawa Ficka*, odnoszącego się do wszelkich procesów dyfuzji.

we wszystkich 3 wymiarach (co jest założeniem bliższym rzeczywistości niż przyjęte w poprzednio rozważanym modelu), to dojdziemy do wzoru identycznego z (10/55). Przy wyznaczaniu współczynnika dyfuzji D_n należy jednakże uwzględnić nie tylko wartości, lecz także kierunki prędkości ruchów termicznych elektronów. Prowadzi to do nieco innej postaci wzoru na D_n, tj.

$$D_n = \frac{1}{3} v_{th}^2 \tau \tag{10/56}$$

Opierając się na zależności (10/9), można wzór ten przedstawić w postaci

$$D_n = \frac{q\tau}{m_n^*} \frac{kT}{q} \tag{10/57}$$

Pierwszy czynnik w tym wyrażeniu jest równy ruchliwości elektronów μ_n (zob. wzór (10/28)), drugi zaś reprezentuje napięcie, pod wpływem którego elektron uzyskuje energię kinetyczną równą kT. Wielkość tę, występującą w wielu wzorach, nazywamy *napięciem termicznym* i oznaczamy symbolem U_T. Można traktować ją jako odpowiednik temperatury wyrażony w woltach zgodnie z zależnością

$$U_T = \frac{kT}{q} = 8{,}62 \cdot 10^{-5}\, T \tag{10/58}$$

W temperaturze pokojowej (tzn. $T = 300$ K) U_T wynosi w przybliżeniu 26 mV. Uwzględniając podane tu związki, można wzór (10/57) napisać w postaci

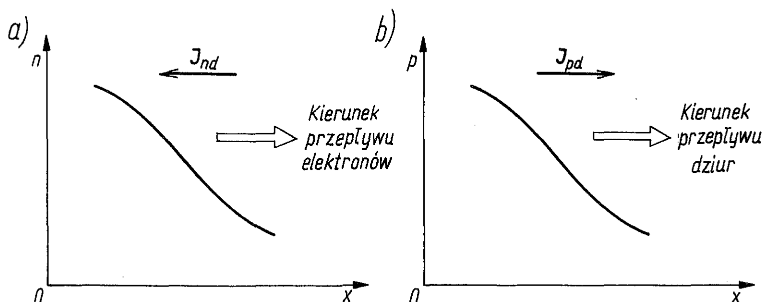
$$\boxed{D_n = \mu_n U_T} \tag{10/59}$$

Równanie to, i analogiczna zależność dla dziur, noszą nazwę *relacji Einsteina*. Wiążą one ze sobą 2 ważne parametry materiałowe półprzewodnika: ruchliwość i współczynnik dyfuzji danego rodzaju nośników. Równanie to zostało tu wyprowadzone w sposób przybliżony. Uważny Czytelnik dostrzegł niewątpliwie pewne nieścisłości w przedstawionych wywodach. W następnym rozdziale wykażemy jednak w sposób bardziej ścisły, że równanie (10/59) jest słuszne dla dowolnego półprzewodnika niezdegenerowanego, a w odniesieniu do nośników mniejszościowych – także dla półprzewodnika zdegenerowanego.

Ze wzoru (10/59) wynika w sposób oczywisty zależność współczynnika dyfuzji D_n od rodzaju półprzewodnika, koncentracji domieszek i temperatury.

W punkcie 10.3 wykazaliśmy (zob. wzór (10/16)), że gęstość prądu elektronowego jest równa iloczynowi koncentracji, ładunku i prędkości tych nośników. Znając gęstość prądu dyfuzyjnego możemy zatem wyznaczyć średnią prędkość dyfuzji elektronów. Wynosi ona

$$\boxed{v_{nd} = -\frac{D_n}{n} \frac{dn}{dx}} \tag{10/60}$$



10/20 Kierunek prądu wywołanego dyfuzją elektronów (a) i dziur (b)

Analogiczne rozumowanie prowadzi do następujących wzorów dotyczących dyfuzji dziur:

$$J_{pd} = -qD_p \frac{\mathrm{d}p}{\mathrm{d}x} \tag{10/61}$$

$$D_p = \mu_p U_T \tag{10/62}$$

$$v_{pd} = -\frac{D_p}{p} \frac{\mathrm{d}p}{\mathrm{d}x} \tag{10/63}$$

Różnica znaku we wzorach (10/61) i (10/55) wynika z różnicy znaków ładunków dziur i elektronów. Jak pokazano na rys. 10/20, zarówno elektrony, jak i dziury poruszają się w kierunku ich malejących koncentracji. Prąd dziurowy płynie jednakże w przeciwnym kierunku niż prąd elektronowy.

Rozważmy teraz bardziej ogólny przypadek, gdy obok nierównomiernego rozkładu koncentracji nośników występuje pole elektryczne. Jeżeli prędkości ruchów uporządkowanych są dostatecznie małe w stosunku do prędkości termicznych, to jak już zaznaczono, procesy unoszenia i dyfuzji nośników zachodzą niezależnie od siebie. W tych warunkach gęstości prądu elektronowego i prądu dziurowego przy przepływie jednowymiarowym wynoszą odpowiednio:

$$J_n = qn\mu_n E + qD_n \frac{\mathrm{d}n}{\mathrm{d}x} \tag{10/64}$$

$$J_p = qp\mu_p E - qD_p \frac{\mathrm{d}p}{\mathrm{d}x} \tag{10/65}$$

Korzystając z relacji Einsteina, można ze wzorów tych wyeliminować współczynniki dyfuzji. Otrzymuje się wówczas:

$$J_n = q\mu_n \left(nE + U_T \frac{\mathrm{d}n}{\mathrm{d}x} \right) \tag{10/66}$$

$$J_p = q\mu_p \left(pE - U_T \frac{\mathrm{d}p}{\mathrm{d}x} \right) \tag{10/67}$$

Ogólniejsze wzory, słuszne dla przepływów trójwymiarowych, mają postać

$$J_n = -qn\,\mu_n \nabla V + qD_n \nabla n \qquad (10/68)$$

$$J_p = -qp\,\mu_p \nabla V - qD_p \nabla p \qquad (10/69)$$

przy czym natężenie pola E zastąpiliśmy tu gradientem potencjału zgodnie z zależnością

$$E = -\nabla V \qquad (10/70)$$

Prądy J_n i J_p zależą więc od gradientu potencjału elektrycznego i od gradientu koncentracji nośników. Gdy ponadto uwzględni się nierównomierny rozkład temperatury w próbce półprzewodnika, to pojawi się jeszcze jeden składnik prądu – proporcjonalny do gradientu temperatury. Zagadnienia te wykraczają jednakże poza zakres tej książki.

11

ŁADUNEK PRZESTRZENNY W PÓŁPRZEWODNIKU

W rozdziale 8 rozpatrywaliśmy półprzewodnik jednorodny w stanie równowagi termodynamicznej. Stwierdziliśmy tam, że w półprzewodniku takim występują w ogólnym przypadku 4 rodzaje ładunków elektrycznych (zjonizowane donory i akceptory oraz elektrony i dziury), przy czym rozkład ich koncentracji jest równomierny w całej objętości półprzewodnika. W tych warunkach koncentracje ładunków dodatnich i ujemnych są jednakowe, a więc ładunki te kompensują się wzajemnie. Znajduje to swój wyraz w warunku *neutralności elektrycznej* wyrażonym wzorem (8/37). W półprzewodniku jednorodnym, ze względu na kompensację ładunków elektrycznych, nie występuje pole elektryczne, a więc potencjał elektryczny ma w każdym punkcie półprzewodnika jednakową wartość.

W półprzewodnikach niejednorodnych, a także w warunkach odbiegających od równowagi termodynamicznej, mamy często do czynienia z sytuacją, w której wzajemna kompensacja ładunków dodatnich i ujemnych nie zachodzi, albo też zachodzi tylko w przybliżeniu. W wyniku tego pozostaje pewien nieskompensowany ładunek rozłożony w objętości półprzewodnika. Nosi on nazwę *ładunku przestrzennego* (lub ładunku objętościowego). Ładunek ten odgrywa ważną rolę przy analizie wielu zjawisk zachodzących w półprzewodnikach, jest on bowiem źródłem dodatkowego pola elektrycznego w pewnych obszarach półprzewodnika.

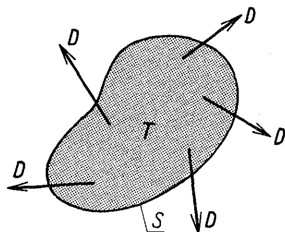
W rozdziale tym zapoznamy się z tzw. *równaniem Poissona*, uzależniającym rozkład pola elektrycznego w półprzewodniku od rozkładu ładunku przestrzennego, a następnie, opierając się na tym równaniu, rozpatrzymy zjawiska występujące w półprzewodniku niejednorodnym w stanie równowagi termodynamicznej.

11.1 RÓWNANIE POISSONA

Jednym z podstawowych praw elektrostatyki, znanym Czytelnikowi ze szkolnego kursu fizyki, jest *prawo Gaussa*, które orzeka, że strumień wektora indukcji elektrycznej $\boldsymbol{D}$ przechodzący przez zamkniętą powierzchnię S jest równy całkowitemu ładunkowi Q zawartemu w obszarze T ograniczonym tą powierzchnią – rys. 11/1. Prawo to można wyrazić wzorem

$$\oint_S D_n \, d\sigma = Q \qquad (11/1)$$

gdzie: D_n – składowa normalna wektora $\boldsymbol{D}$; dσ – element pola powierzchni S.



11/1
Wielkości występujące we wzorze Gaussa

Jeżeli ładunek Q jest rozłożony w obszarze T, w taki sposób, że jego gęstość objętościowa

$$\varrho = \frac{dQ}{d\tau} \qquad (11/2)$$

ma wszędzie wartość skończoną, to obie strony równania (11/1) można podzielić przez objętość tego obszaru, a następnie przejść do granicy, zmniejszając do zera średnicę obszaru. Po lewej stronie równania otrzymamy wówczas *dywergencję* wektora $\boldsymbol{D}$[1], po prawej zaś – wspomnianą gęstość ładunku przestrzennego ϱ. W ten sposób dochodzimy do równania

$$\nabla \cdot \boldsymbol{D} = \varrho \qquad (11/3)$$

które jest jednym z 4 równań Maxwella – zob. dodatek A.

Jeżeli rozpatrywany półprzewodnik jest izotropowy ze względu na przenikalność elektryczną ε, co najczęściej występuje, to

$$\boldsymbol{D} = \varepsilon \boldsymbol{E} \qquad (11/4)$$

gdzie $\boldsymbol{E}$ – natężenie pola elektrycznego.

---

[1] *Dywergencją* pola wektorowego $\boldsymbol{D}$ nazywamy skalar, oznaczany symbolem div $\boldsymbol{D}$ lub $\nabla \cdot \boldsymbol{D}$, określony w następujący sposób:

$$\nabla \cdot \boldsymbol{D} = \lim_{V \to 0} \frac{\oint_S D_n \, d\sigma}{V}$$

gdzie V – objętość obszaru ograniczonego zamkniętą powierzchnią S. W układzie współrzędnych prostokątnych x, y, z dywergencja wektora $\boldsymbol{D}$ wyraża się wzorem

$$\nabla \cdot \boldsymbol{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}$$

Podstawiając (11/4) do (11/3) i uwzględniając zależność (10/70), otrzymamy następujące równanie, zwane *równaniem Poissona*:

$$\nabla^2 V = -\frac{\varrho}{\varepsilon} \tag{11/5}$$

Wyrażenie występujące po lewej stronie równania jest *laplasjanem* potencjału elektrostatycznego V[1].

Równanie Poissona wiąże, jak widzimy, rozkład potencjału z rozkładem ładunku przestrzennego w danym obszarze. Jeżeli w każdym punkcie tego obszaru $\varrho = 0$, to, jak wynika ze wzorów (11/3) i (11/4), dywergencją wektora $\boldsymbol{E}$ jest wszędzie równa zeru. Nie oznacza to jednak, że w półprzewodniku nie występuje pole elektryczne. Pole takie może istnieć, z warunku $\nabla \cdot \boldsymbol{E} = 0$ wynika jednak, że ładunki będące źródłem tego pola znajdują się poza rozpatrywanym obszarem półprzewodnika. Z sytuacją taką mieliśmy do czynienia w p. 10.3 – zob. rys. 10/4. Ładunki będące źródłem pola w półprzewodniku znajdują się w tym przypadku na powierzchniach kontaktów.

We współrzędnych prostokątnych równanie (11/5) przyjmuje postać:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\varrho(x, y, z)}{\varepsilon} \tag{11/6}$$

W dalszych punktach tej książki ograniczymy się do zagadnień jednowymiarowych. Zgodnie z tym założymy więc, że zarówno potencjał, jak i gęstość ładunku przestrzennego zależą jedynie od współrzędnych x. Równanie Poissona upraszcza się wówczas do jednej z następujących równoważnych postaci:

$$\frac{\mathrm{d}^2 V}{\mathrm{d}x^2} = -\frac{\varrho(x)}{\varepsilon} \tag{11/7a}$$

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{\varrho(x)}{\varepsilon} \tag{11/7b}$$

Wartości przenikalności elektrycznej ważniejszych półprzewodników podano w dodatku G. Dla krzemu $\varepsilon = 11{,}9\ \varepsilon_0 = 1{,}05 \cdot 10^{-12}$ F/cm.

Biorąc pod uwagę wszystkie 4 rodzaje ładunków występujących w półprzewodniku, można napisać następujący wzór na gęstość ładunku przestrzennego:

$$\varrho = q\,(N_d^+ - N_a^- + p - n) \tag{11/8}$$

Gdyby wielkości występujące w tym wzorze były znanymi funkcjami współrzędnej x, to na podstawie równania (11/7) moglibyśmy wyznaczyć rozkład potencjału i natężenia pola w półprzewodniku. Tak niestety nie jest. Rozkład ładunków elektrycznych (przede wszystkim elektronów i dziur, a w pewnych

[1] *Laplasjan* (operator Laplace'a) jest to operator różniczkowy przyporządkowujący funkcji skalarnej V dywergencję jej gradientu

$$\nabla^2 V = \operatorname{div} \operatorname{grad} V$$

warunkach także zjonizowanych domieszek[1]) zależy bowiem od rozkładu potencjału w półprzewodniku, a jednocześnie przebieg potencjału, poprzez równanie Poissona, jest uzależniony od rozkładu gęstości tych ładunków. Jet to przykład występującego często w fizyce tzw. *zagadnienia samouzgodnionego*. Rozpatrywany tu problem nie daje się rozwiązać w sposób ogólny. Można jedynie znaleźć rozwiązania przybliżone dla określonych przypadków.

11.2 PÓŁPRZEWODNIK JEDNORODNY

Dla półprzewodnika jednorodnego koncentracje elektronów i dziur można wyrazić za pomocą wzorów (9/1) i (9/2). Podstawiając te wyrażenia do (11/8) i zakładając, że wszystkie domieszki są zjonizowane, otrzymujemy

$$\varrho = q\,(N_d - N_a + p_0 + \Delta p - n_0 - \Delta n) \qquad (11/9)$$

gdzie: n_0 i p_0 – koncentracje równowagowe, a Δn i Δp – odpowiednie koncentracje nadmiarowe.

Korzystając z warunku neutralności elektrycznej wyrażonego wzorem (8/37) i podstawiając (11/9) do (11/7), otrzymujemy następujące dwie równoważne postaci jednowymiarowego równania Poissona

$$\frac{d^2V}{dx^2} = \frac{q}{\varepsilon}(\Delta n - \Delta p) \qquad (11/10a)$$

$$\frac{dE}{dx} = \frac{q}{\varepsilon}(\Delta p - \Delta n) \qquad (11/10b)$$

Ze wzorów tych będziemy korzystali w następnym rozdziale przy analizowaniu transportu nośników nadmiarowych w jednorodnym półprzewodniku

11.3 PÓŁPRZEWODNIK NIEJEDNORODNY

W większości elementów półprzewodnikowych występuje nierównomierny rozkład koncentracji domieszek. Można nawet powiedzieć, że działanie większości elementów półprzewodnikowych jest możliwe dzięki temu, że do monokrystalicznej próbki półprzewodnika potrafimy wprowadzać domieszki donorowe i akceptorowe tak, aby uzyskać ich pożądane rozkłady.

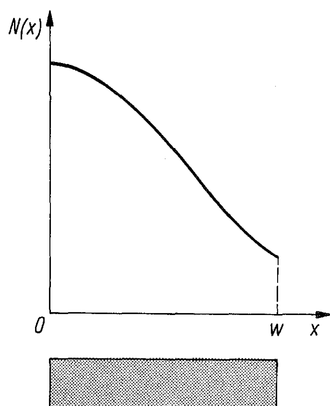
---

[1]) W rozpatrywanych tu zagadnieniach przyjmujemy, że koncentracje domieszek są znanymi a priori a funkcjami współrzędnej x. Stopień jonizacji tych domieszek zależy jednakże od odległości poziomu Fermiego od odpowiednich poziomów domieszkowych. Odległości te zależą z kolei od potencjału w danym punkcie półprzewodnika. W wyniku tego rozkład $N_d^+(x)$ i $N_a^-(x)$ może być w pewnych warunkach uzależniony od rozkładu potencjału w półprzewodniku. W dalszych rozważaniach przyjmiemy jednakże dla uproszczenia, że domieszki donorowe i akceptorowe są całkowicie zjonizowane, tzn. $N_d^+ = N_d$ oraz $N_a^- = N_a$.

Największe znaczenie praktyczne mają tzw. *złącza p-n*, tzn. struktury półprzewodnikowe, w których koncentracje donorów i akceptorów zmieniają się w taki sposób, że w pewnym przekroju następuje zmiana typu przewodnictwa z elektronowego na dziurowe. W strukturze takiej można więc wyodrębnić 2 sąsiadujące ze sobą obszary, z których jeden wykazuje przewodnictwo typu p, drugi zaś – typu n. Tym tłumaczy się nazwa złącze p-n.

W książce tej nie będziemy analizowali zjawisk występujących w złączu p-n, zajmiemy się natomiast znacznie prostszym przypadkiem półprzewodnika niejednorodnego, który w całej swej objętości zachowuje ten sam typ przewodnictwa. Innymi słowy przedmiotem naszych rozważań będzie półprzewodnik domieszkowany niejednorodnie, w którym nie występuje złącze p-n. Rozważania nasze będą dotyczyły stanu równowagi termodynamicznej.

Weźmy pod uwagę prostopadłościenną próbkę półprzewodnika, w której efektywna koncentracja domieszek, zdefiniowana wzorem (8/40), zmienia się w funkcji współrzędnej x w sposób przedstawiony na rys. 11/2. Rozkład koncentracji domieszek o takim charakterze otrzymuje się np. przy domieszkowaniu półprzewodnika metodą dyfuzji atomów domieszkowych w wysokich temperaturach. Funkcja $N(x)$ opisuje tzw. *profil domieszkowania* półprzewodnika.



11/2
Rozkład koncentracji domieszek w półprzewodniku niejednorodnym

Ze względu na to, że w omawianym przypadku koncentracja $N(x)$ jest wszędzie dodatnia, mamy tu do czynienia z półprzewodnikiem typu n. W celu ułatwienia dyskusji załóżmy, że w półprzewodniku tym występują jedynie domieszki donorowe. Wyniki naszych rozważań i wyprowadzone wzory będą jednakże słuszne także w przypadku, gdy obok donorów występują akceptory, z tym zastrzeżeniem, że wówczas $N(x) = N_d(x) - N_a(x) > 0$.

Załóżmy ponadto, że rozpatrywany półprzewodnik jest dostatecznie silnie domieszkowany, tzn. że w każdym punkcie koncentracja $N(x)$ jest znacznie większa niż koncentracja samoistna n_i. Przy tych założeniach można w wyrażeniu (11/8) pominąć koncentrację dziur jako znikomo małą wobec koncentra-

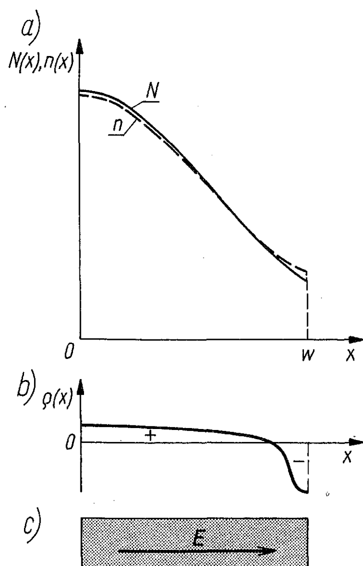
cji elektronów. Dzięki temu równanie Poissona przyjmuje postać

$$\frac{dE}{dx} = \frac{q}{\varepsilon}[N(x)-n(x)] \tag{11/11}$$

Gdyby w każdym punkcie półprzewodnika koncentracja elektronów $n(x)$ była równa koncentracji donorów $N(x)$, tak jak to występuje w półprzewodniku jednorodnym, to ładunki te kompensowałyby się wzajemnie i w półprzewodniku nie występowałoby pole elektryczne. Taki rozkład koncentracji elektronów nie odpowiada jednakże równowadze termodynamicznej, ponieważ występujący w tych warunkach gradient koncentracji tych nośników wywołuje ich dyfuzję z obszarów o większej koncentracji domieszek do obszarów słabiej domieszkowanych.

W analizowanym przypadku (rys. 11/2) elektrony dyfundują w kierunku zgodnym z kierunkiem osi x. W obszarach silniej domieszkowanych koncentracja ich zmaleje, wskutek czego pozostanie tam nieskompensowany ładunek dodatnio zjonizowanych donorów. Na przeciwległym końcu próbki wytworzy się natomiast pewien nadmiar elektronów, w wyniku czego pojawi się tam dodatkowy ładunek ujemny.

Całkowity ładunek dodatni jest oczywiście równy całkowitemu ładunkowi ujemnemu, ponieważ w próbce nie pojawiły się, ani też z niej nie zniknęły, żadne ładunki. Nastąpiło jedynie przemieszczenie części elektronów w kierunku malejącej koncentracji donorów. W wyniku tego przemieszczenia pojawia się w półprzewodniku ładunek przestrzenny (zob. rys. 11/3), a wraz z nim pole elektryczne skierowane zgodnie z kierunkiem osi x. Pole to nazywamy *polem wewnętrznym*, ponieważ jego źródłem są ładunki znajdujące się wewnątrz rozpatrywanego obszaru, albo też *polem wbudowanym*, ze względu na to, że zostało



11/3
Półprzewodnik niejednorodny typu n:
a) rozkład koncentracji donorów i elektronów; b) rozkład ładunku przestrzennego; c) pole wbudowane

ono niejako „wbudowane" w półprzewodnik w wyniku odpowiedniego rozkładu wprowadzonych do niego domieszek.

Siły tego pola działają na elektrony w kierunku przeciwnym do kierunku ich przepływu dyfuzyjnego, przeciwdziałają więc dalszemu przemieszczaniu się elektronów.

Stan równowagi termodynamicznej zostaje osiągnięty, gdy tendencja elektronów do dyfuzji zostanie dokładnie zrównoważona tendencją do przepływu w przeciwnym kierunku pod działaniem sił pola elektrycznego. Równowaga ta następuje więc przy takim rozkładzie ładunku przestrzennego, przy którym w każdym punkcie półprzewodnika gęstość prądu elektronowego staje się równa zeru.

Korzystając ze wzoru (10/66), możemy napisać, że dla półprzewodnika typu n w stanie równowagi termodynamicznej

$$q\,\mu_n\left(nE + U_T\frac{\mathrm{d}n}{\mathrm{d}x}\right) = 0 \tag{11/12}$$

Z równania tego wynika następujący wzór na natężenie pola elektrycznego:

$$E = -\frac{U_T}{n}\frac{\mathrm{d}n}{\mathrm{d}x} = -U_T\frac{\mathrm{d}}{\mathrm{d}x}(\ln n) \tag{11/13}$$

Do wyznaczenia natężenia pola jest więc konieczna znajomość rozkładu koncentracji elektronów $n(x)$.

Na podstawie rys. 11/3 możemy stwierdzić, że rozkład $n(x)$ nie pokrywa się dokładnie z rozkładem koncentracji donorów $N(x)$. W wielu sytuacjach (jak dalej zobaczymy, przy dostatecznie dużych koncentracjach donorów i niezbyt szybkich zmianach tych koncentracji w funkcji współrzędnej x) można jednakże przyjmować, że rozkłady te różnią się bardzo nieznacznie, tzn., że

$$n(x) \approx N(x) \tag{11/14}$$

Z faktu tego nie można jednak wyciągnąć wniosku, że różnica $N(x) - n(x)$, występująca po prawej stronie równania (11/11) jest pomijalnie mała. Z równania tego wynika bowiem, że to właśnie ta różnica jest odpowiedzialna za pole elektryczne, które wytwarza się w półprzewodniku. Ładunek przestrzenny o gęstości $\varrho(x)$, będący źródłem tego pola, może być bardzo mały w stosunku do ładunku donorów, którego gęstość wynosi $qN(x)$, ale pole wytworzone przez ten ładunek ma istotne znaczenie[1].

Przybliżona równość (11/14) oznacza, że ładunki donorów i elektronów kompensują się niemal całkowicie. Stan taki określamy jako *quasi-neutralność* elektryczną. Założenie quasi-neutralności będzie podstawą dalszych rozważań nie tylko w warunkach równowagi termodynamicznej, lecz również przy rozpatrywaniu transportu nośników nadmiarowych. W punkcie 11.5 spróbujemy zbadać, w jakich sytuacjach wolno jest korzystać z tego założenia.

[1] Ogólnie z faktu, że 2 liczby są w przybliżeniu równe, nie można wyciągać wniosku, że ich różnica jest do pominięcia. Przykładowo, okręt bez kapitana waży prawie tyle samo, co z kapitanem. Nie oznacza to jednak, że kapitan nic nie waży.

Posługując się przybliżeniem (11/14), należy jednak zawsze postępować z dużą ostrożnością. Podstawiając np. $n(x) \approx N(x)$ do wzoru (11/13), należy uświadamiać sobie, że nawet gdy same funkcje $n(x)$ i $N(x)$ są niemal równe, ich gradienty mogą różnić się dość znacznie.

Mając na uwadze to zastrzeżenie, można na podstawie wzorów (11/13) i (11/14) napisać

$$E(x) \approx -\frac{U_T}{N(x)} \frac{\mathrm{d}N}{\mathrm{d}x} = -U_T \frac{\mathrm{d}}{\mathrm{d}x} \ln N(x) \tag{11/15}$$

Ponieważ $N(x)$ jest znane, zatem otrzymany wzór umożliwia wyznaczenie *przybliżonej* wartości natężenia pola w dowolnym punkcie x.

Analogiczne rozumowanie dla półprzewodnika typu p prowadzi do wzoru

$$E = \frac{U_T}{p} \frac{\mathrm{d}p}{\mathrm{d}x} = U_T \frac{\mathrm{d}}{\mathrm{d}x} \ln p \tag{11/16}$$

Korzystając z założenia quasi-neutralności

$$p(x) \approx |N(x)| \tag{11/17}$$

otrzymujemy

$$E(x) \approx \frac{U_T}{N(x)} \frac{\mathrm{d}N}{\mathrm{d}x} = U_T \frac{\mathrm{d}}{\mathrm{d}x} \ln |N(x)| \tag{11/18}$$

przy czym $|N(x)|$ oznacza w tym przypadku koncentrację akceptorów w punkcie x.

11.4 MODEL PASMOWY PÓŁPRZEWODNIKA NIEJEDNORODNIE DOMIESZKOWANEGO

W celu narysowania modelu pasmowego półprzewodnika rozpatrywanego w poprzednim punkcie posłużymy się wzorem (8/47), który tu z pewnymi modyfikacjami powtórzymy

$$W_F - W_i(x) = kT \ln \frac{n(x)}{n_i} \tag{11/19}$$

Wzór (8/47) odnosi się do niezdegenerowanego półprzewodnika jednorodnego, w którym koncentrację elektronów, niezależną od współrzędnej x, oznaczaliśmy symbolem n_0. W analizowanym półprzewodniku niejednorodnym koncentracja ta zmienia się wzdłuż osi x. W wyniku tego zmienia się również położenie poziomu samoistnego W_i.

Wartość energii Fermiego W_F jest natomiast niezależna od x. Jest to ogólne prawo dotyczące stanu równowagi termodynamicznej dowolnego układu połączonych ze sobą elektrycznie przewodników. Układ taki musi charakteryzować się stałym poziomem Fermiego, ponieważ poziom ten określa, jak wiadomo, prawdopodobieństwo obsadzenia stanów kwantowych przez elektrony. Gdyby w różnych częściach układu energia Fermiego miała różne wartości,

to oznaczałoby to, że prawdopodobieństwo obsadzenia stanów o danej energii W jest w pewnych częściach układu większe niż w innych. Nastąpiłby zatem przepływ elektronów z obszarów o większym do obszarów o mniejszym prawdopodobieństwie obsadzenia. Z definicji stanu równowagi termodynamicznej wynika natomiast, że w stanie tym nie występują żadne przepływy.

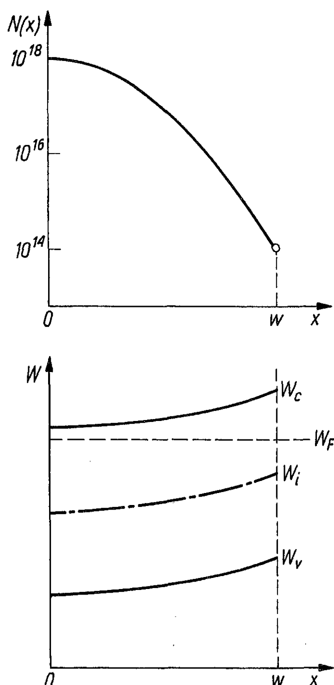
W układzie, który początkowo nie znajdował się w równowadze termodynamicznej, nastąpi przepływ nośników w takim kierunku i w takiej ilości, aż ustali się wspólny dla całego układu poziom Fermiego. Z tą chwilą wszelkie przepływy ustaną.

Zjawisko ustalania się jednakowego poziomu Fermiego w całym układzie jest w pewnym sensie analogiczne do wyrównywania się poziomu cieczy w systemie naczyń połączonych. I tu, jeżeli poziomy te początkowo nie są wyrównane, nastąpią przepływy cieczy, które trwać będą dopóty, aż w całym systemie ustali się jednakowy ich poziom.

Dla półprzewodnika typu n, spełniającego założenia przyjęte w poprzednim punkcie, można, opierając się na zależności (11/14), napisać następujący wzór:

$$W_F - W_i(x) = kT \ln \frac{N(x)}{n_i} \tag{11/20}$$

Ponieważ profil domieszkowania $N(x)$ jest znany, zatem na podstawie wzoru (11/20) można wykreślić przedstawiony na rys. 11/4 model pasmowy rozpatrywanego półprzewodnika.



11/4
Półprzewodnik niejednorodny typu *n*: a) rozkład koncentracji donorów; b) model pasmowy

Na rysunku tym można zauważyć, że w miarę przesuwania się wzdłuż osi x poziom Fermiego oddala się coraz bardziej od dna pasma przewodnictwa. Jest to związane ze zmniejszaniem się koncentracji elektronów wzdłuż próbki półprzewodnika (rys. 11/4a). Pochylenie pasm energetycznych świadczy o obecności pola elektrycznego w próbce. Natężenie tego pola można wyznaczyć ze wzoru (10/15), który tu dla wygody powtórzymy

$$E = \frac{1}{q} \frac{dW_i}{dx} \tag{11/21}$$

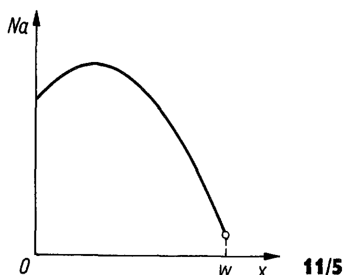
Ponieważ pochodna dW_i/dx jest dodatnia, zatem kierunek tego pola jest zgodny z kierunkiem osi x, tak jak to pokazano na rys. 11/3.

Z kolei zakrzywienie pasm energetycznych dowodzi obecności ładunku przestrzennego w półprzewodniku. Ze wzorów (11/7b) i (11/21) otrzymujemy bowiem

$$\varrho(x) = \frac{\varepsilon}{q} \frac{d^2 W_i}{dx^2} \tag{11/22}$$

Znak pochodnej d^2W_i/dx^2 jest w całej próbce dodatni, co świadczyłoby o tym, że gęstość ładunku przestrzennego jest wszędzie dodatnia. Jest to niezgodne z rzeczywistością (zob. rys. 11/3), a przyczyną tego jest przybliżony charakter naszych rozważań. Opieramy się w nich bowiem na przybliżonej równości (11/14) i chociaż niedokładność wzoru (11/20) jest znikomo mała, to jednak przy wyznaczaniu z niego drugiej pochodnej funkcji $W_i(x)$ dochodzimy do wyników znacznie odbiegających od rzeczywistości.

Problem 11.1 Dla półprzewodnika typu p o profilu domieszkowania podanym na rys. 11/5 narysuj model pasmowy oraz przebieg natężenia pola elektrycznego i gęstości ładunku przestrzennego w funkcji x.



11/5

11.5* ZAŁOŻENIE QUASI-NEUTRALNOŚCI ELEKTRYCZNEJ

Założenie (11/14) i oparte na nim wywody są tym dokładniejsze im różnica $N-n$ jest mniejsza w stosunku do N w rozważanym obszarze półprzewodnika. Tak więc dla półprzewodnika typu n jako miarę odchylenia od warunku neutralności elektrycznej można przyjąć wielkość

$$\delta = \left| \frac{N(x)-n(x)}{N(x)} \right| \tag{11/23}$$

W celu wyznaczenia tego współczynnika skorzystamy z równania (11/11), które po odpowiednim przekształceniu przyjmuje postać

$$\frac{N(x)-n(x)}{N(x)} = \frac{\varepsilon}{qN(x)} \frac{\mathrm{d}E}{\mathrm{d}x} \tag{11/24}$$

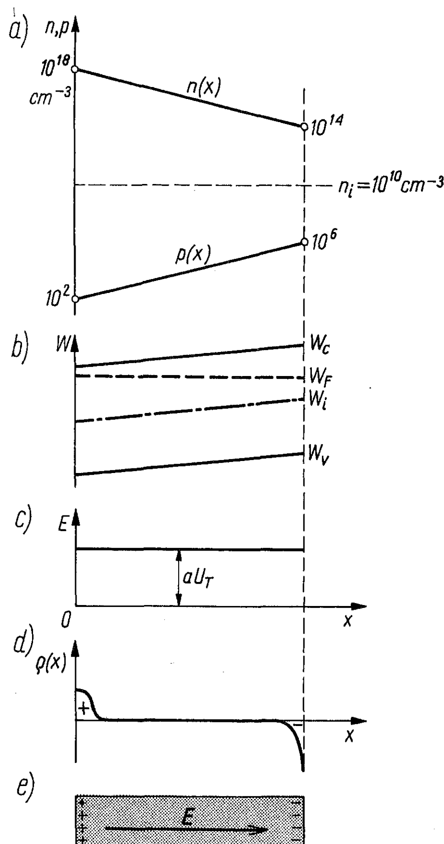
Przybliżoną wartość pochodnej dE/dx możemy wyznaczyć ze wzoru (11/15). Po podstawieniu do (11/24) otrzymamy wówczas

$$\frac{N(x)-n(x)}{N(x)} \approx \frac{\varepsilon U_T}{q\,N(x)} \frac{\mathrm{d}^2}{\mathrm{d}x^2} \ln N(x) \tag{11/25}$$

Prawa strona tego równania staje się równa zeru, jeżeli ln $N(x)$ jest liniową funkcją x, tzn. gdy rozkład koncentracji donorów jest opisany funkcją wykładniczą

$$N(x) = N(0) \exp(-ax) \tag{11/26}$$

Model pasmowy półprzewodnika o takim profilu domieszkowania przedstawiono na rys. 11/6. Pokazano tam również rozkład natężenia pola elektrycznego, gęstości ładunku przestrzennego oraz koncentracji nośników większościowych i mniejszościowych. Tych ostatnich nie braliśmy pod uwagę w rozważaniach prowadzonych w p. 11.3 i 11.4.



11/6 Półprzewodnik o wykładniczym rozkładzie donorów: a) rozkład koncentracji nośników; b) model pasmowy; c) rozkład natężenia pola elektrycznego; d) rozkład ładunku przestrzennego; e) pole wbudowane i jego źródła

Natężenie pola elektrycznego można wyznaczyć ze wzoru (11/15). Wynosi ono

$$E = a\, U_T = \text{const} \tag{11/27}$$

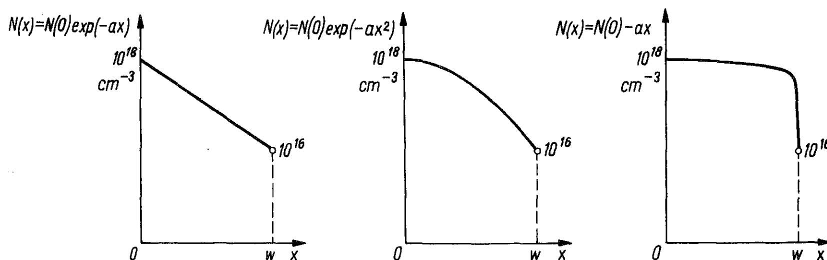
Pole elektryczne w badanym półprzewodniku jest więc jednorodne, a gęstość ładunku przestrzennego jest prawie wszędzie równa zeru. Źródłem tego pola są ładunki zgromadzone w bardzo cienkich warstwach przypowierzchniowych w przeciwległych końcach próbki – zob. rys. 11/6. Poza tymi warstwami zachodzi całkowita kompensacja ładunków dodatnich i ujemnych, a więc są tu dokładnie spełnione warunki neutralności elektrycznej.

Przy innych profilach domieszkowania kompensacja ta nie jest całkowita, a założenie quasi-neutralności elektrycznej jest spełnione tym gorzej, im większą wartość ma wyrażenie

$$\frac{1}{N(x)}\,\frac{d^2}{dx^2}\ln N(x)$$

a więc im słabiej jest domieszkowany półprzewodnik i im bardziej funkcja $\ln N(x)$ odbiega od liniowej.

Na rysunku 11/7 zestawiono wykresy $\ln N(x)$ dla 3 różnych profili domieszkowania. Na podstawie tego rysunku można stwierdzić, że z podanych tu przykładów warunek quasi-neutralności jest spełniony najgorzej dla liniowego rozkładu koncentracji domieszek.



11/7 Różne rozkłady koncentracji domieszek: a) rozkład wykładniczy; b) rozkład Gaussa; c) rozkład liniowy

Założenie quasi-neutralności elektrycznej zawodzi natomiast całkowicie, gdy efektywna koncentracja domieszek zmienia znak, tzn. w złączu *p-n*.

Problem 11.2. Oszacuj względne odchylenie od neutralności elektrycznej dla rozkładu koncentracji domieszek opisanego funkcją Gaussa

$$N(x) = N(0)\exp(-ax^2) \tag{11/28}$$

i przedstawionego na rys. 11/7b. Przyjmijmy, że długość próbki $w = 2$ μm.
Rozwiązanie: Dla tego profilu domieszkowania prawa strona równania (11/25) przyjmuje postać

$$\frac{\varepsilon U_T}{qN(x)}\,\frac{d^2}{dx^2}\ln N(x) = -2a\,\frac{\varepsilon U_T}{qN(x)} \tag{11/29}$$

Osiąga ona największą wartość w punkcie $x = w$.

Ze wzoru (11/28) wynika, że

$$a = \frac{1}{w^2} \ln \frac{N(0)}{N(w)} \tag{11/30}$$

a zatem maksymalna wartość współczynnika δ wynosi

$$\delta_{max} \approx \frac{2}{w^2} \frac{\varepsilon U_T}{qN(w)} \ln \frac{N(0)}{N(w)} \tag{11/31}$$

Podstawiając odpowiednie wartości stałych fizycznych oraz $N(0) = 10^{18}$ cm^{-3}, $N(w) = 10^{16}$ cm^{-3}, $w = 2 \cdot 10^{-4}$ cm, znajdziemy, że

$$\delta_{max} \approx 0{,}004$$

Otrzymany wynik wskazywałby na fakt, że warunek quasi-neutralności elektrycznej jest tu spełniony z dobrym przybliżeniem. Należy jednak zdawać sobie sprawę z małej dokładności dokonanej analizy, o czym wspominaliśmy już wcześniej. Korzystamy tu bowiem z drugiej pochodnej funkcji, której przebieg znamy tylko w sposób przybliżony, a przy każdym różniczkowaniu niedokładność może ulec znacznemu powiększeniu.

11.6 RÓWNANIA BOLTZMANNA

Przy analizowaniu zjawisk zachodzących w półprzewodnikach zachodzi konieczność posługiwania się nie tylko natężeniem pola elektrycznego, lecz również potencjałem elektrycznym. Ponieważ potencjał jest określony z dokładnością do stałej addytywnej, zatem jego wartość może być mierzona względem dowolnie obranego poziomu odniesienia. W różnych zagadnieniach dogodnie jest poziom ten ustalać w różny sposób. W przypadkach rozpatrywanych w tym rozdziale jako poziom odniesienia potencjału dogodnie jest obrać poziom Fermiego, jest on bowiem jednakowy w całej próbce półprzewodnika, który zgodnie z przyjętymi założeniami znajduje się w stanie równowagi termodynamicznej.

Jako definicję potencjału w półprzewodniku przyjmiemy zatem wzór

$$\boxed{\varphi(x) = \frac{W_i(x) - W_F}{-q} = \frac{1}{q}\left[W_F - W_i(x)\right]} \tag{11/32}$$

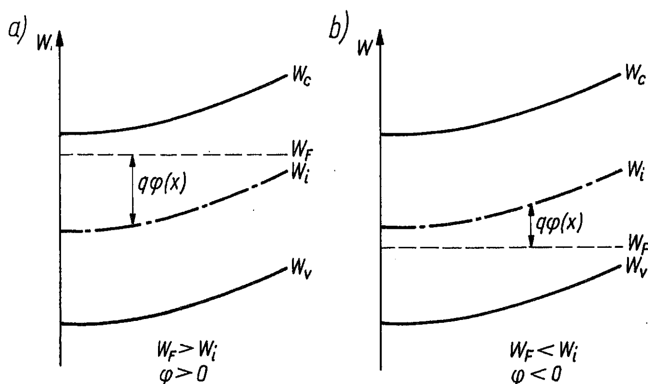
Jak łatwo można sprawdzić, ze wzoru tego otrzymuje się przytoczone już poprzednio wyrażenie na natężenie pola elektrycznego – wzór (11/21).

Tak zdefiniowany potencjał jest dodatni w półprzewodniku typu *n*, ujemny zaś w półprzewodniku typu *p* – zob. rys. 11/8. W przypadku złącza *p-n* przyjmuje on wartość równą zeru w przekroju, w którym $n(x) = p(x) = n_i$.

Korzystając z zależności (11/32), można wzór (11/19) przekształcić do postaci

$$\boxed{n(x) = n_i \exp \frac{\varphi(x)}{U_T}} \tag{11/33}$$

Znając rozkład potencjału w półprzewodniku, można więc wyznaczyć rozkład koncentracji elektronów.



11/8 Potencjał w półprzewodniku typu *n* (a) i typu *p* (b)

Wzór (11/33) jest słuszny dla dowolnego niezdegenerowanego półprzewodnika w warunkach równowagi termodynamicznej. Dla takiego półprzewodnika o dowolnym rozkładzie koncentracji domieszek możemy zastosować prawo działania mas, zgodnie z którym

$$n(x)\,p(x) = n_i^2 \tag{11/34}$$

Uwzględniając to, otrzymujemy następujący wzór opisujący rozkład koncentracji dziur:

$$\boxed{p\,(x) = n_i \exp \frac{-\varphi(x)}{U_T}} \tag{11/35}$$

Równania (11/33) i (11/35) noszą nazwę *równań Boltzmanna.* Korzystając z nich, można jednowymiarowe równanie Poissona przedstawić w postaci

$$\frac{d^2\varphi}{dx^2} = \frac{q}{\varepsilon}\left[2n_i \operatorname{sh}\frac{\varphi(x)}{U_T} - N(x)\right] \tag{11/36}$$

gdzie $N(x)$ – efektywna koncentracja domieszek. Przy znanym rozkładzie domieszek równanie to umożliwia wyznaczenie rozkładu potencjału w półprzewodniku. Niestety jest to równanie przestępne, które nie może być rozwiązane analitycznie nawet dla najprostszych przypadków niejednorodnie domieszkowanego półprzewodnika. Równanie to można rozwiązać jedynie metodami przybliżonymi, np. numerycznymi.

Opierając się na równaniach Boltzmanna, wykażemy słuszność podanych w poprzednim rozdziale relacji Einsteina. W stanie równowagi termodynamicznej zarówno prąd elektronowy, jak i dziurowy są równe zeru, a zatem:

$$J_n = qn\mu_n E + qD_n \frac{dn}{dx} = 0 \tag{11/37}$$

$$J_p = qp\mu_p E - qD_p \frac{dp}{dx} = 0 \tag{11/38}$$

Podstawiając $E = -\mathrm{d}\varphi/\mathrm{d}x$ i wyznaczając pochodne $\mathrm{d}n/\mathrm{d}x$ i $\mathrm{d}p/\mathrm{d}x$ z równań Boltzmanna, otrzymujemy:

$$-qn\mu_n \frac{\mathrm{d}\varphi}{\mathrm{d}x} + qD_n \frac{n}{U_T} \frac{\mathrm{d}\varphi}{\mathrm{d}x} = 0 \tag{11/39}$$

$$-qp\mu_p \frac{\mathrm{d}\varphi}{\mathrm{d}x} + qD_p \frac{p}{U_T} \frac{\mathrm{d}\varphi}{\mathrm{d}x} = 0 \tag{11/40}$$

a zatem

$$\boxed{D_n = \mu_n U_T} \tag{11/41}$$

$$\boxed{D_p = \mu_p U_T} \tag{11/42}$$

Zależności te, podobnie jak równania Boltzmanna, są słuszne dla półprzewodnika niezdegenerowanego.

Dla półprzewodnika zdegenerowanego słuszny jest wzór dotyczący nośników mniejszościowych, natomiast dla nośników większościowych obowiązuje zależność, która np. dla półprzewodnika typu n, ma następującą postać:

$$D_n = \mu_n U_T \frac{F_{1/2}(\eta)}{F_{-1/2}(\eta)} \tag{11/43}$$

gdzie: $F_{1/2}(\eta)$ – całka Fermiego zdefiniowana wzorem (8/11), a $F_{-1/2}(\eta)$ – analogiczna całka o wskaźniku $-1/2$.
Korzystając z rozwinięć funkcji $F(\eta)$ w szereg potęgowy, można wzór (11/43) doprowadzić do postaci

$$D_n = \mu_n U_T \left[1 + 0{,}354 \left(\frac{n}{N_c}\right) - 9{,}9 \cdot 10^{-3} \left(\frac{n}{N_c}\right)^2 + \ldots\right] \tag{11/44}$$

gdzie N_c – efektywna gęstość stanów w pasmie przewodnictwa.

12

TRANSPORT NOŚNIKÓW NADMIAROWYCH W PÓŁPRZEWODNIKU

Zjawiska transportu nośników w półprzewodniku są to procesy polegające na przepływie tych nośników w wyniku zaburzenia stanu równowagi termodynamicznej. Każdy proces transportu można traktować jako współdziałanie następujących dwóch efektów:

a) *oddziaływania sił zewnętrznych* na nośniki, w wyniku czego są one wyprowadzane ze stanu równowagi;

b) *reakcji ośrodka*, wywołanej oddziaływaniem nośników z ośrodkiem, w wyniku czego układ stara się powrócić do stanu równowagi termodynamicznej.

Powrót układu do stanu równowagi odbywa się za pośrednictwem dwóch rodzajów procesów. Jednym z nich jest *rozpraszanie* nośników zachodzące wskutek ich zderzeń z siecią krystaliczną. Cechą charakterystyczną tego procesu jest to, że liczba uczestniczących w nim nośników nie ulega zmianie.

Drugim rodzajem procesów przywracających równowagę termodynamiczną są zjawiska rekombinacyjno-generacyjne, w wyniku których następuje zmiana ilości nośników.

Z procesami transportu nośników zapoznaliśmy się już w rozdz. 10. Stwierdziliśmy tam, że w warunkach, w których nie występują gradienty koncentracji nośników, proces rozpraszania polega na raptownych i przypadkowych zmianach prędkości nośników – jest to tzw. *relaksacja zderzeniowa*. Jeżeli w półprzewodniku występują gradienty koncentracji nośników nadmiarowych, to obok relaksacji zderzeniowej istotną rolę odgrywa inny proces rozproszeniowy, mianowicie *dyfuzja* nośników.

Z chwilą pojawienia się nośników nadmiarowych zachodzą w półprzewodniku procesy rekombinacyjno-generacyjne, z którymi zapoznaliśmy się w rozdz. 9. Obecnie zajmiemy się najbardziej ogólnym przypadkiem transportu nośników nadmiarowych, w którym występuje jednocześnie unoszenie nośników pod dzia-

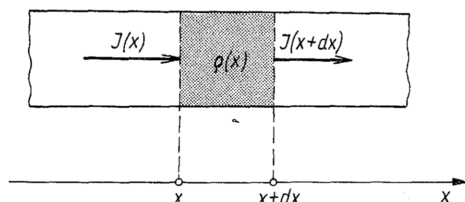
łaniem sił pola elektrycznego, ich dyfuzja, zachodząca wskutek nierównomiernego rozkładu koncentracji, oraz procesy rekombinacyjno-generacyjne.

Rozdział ten stanowi w pewnym sensie podsumowanie i rozszerzenie wiadomości nabytych w poprzednich rozdziałach. Punktem wyjścia rozważań będą równania ciągłości, będące wyrazem zasady zachowania ładunku elektrycznego. Wraz z równaniami prądu i równaniem Poissona tworzą one podstawowy układ równań, na którym opiera się tzw. *fenomenologiczna teoria transportu* sformułowana w 1950 r. przez Roosbroecka. Teoria ta jest słuszna dla znacznej większości materiałów stosowanych w elementach półprzewodnikowych. Zawodzi natomiast w przypadkach, w których czas życia nośników nadmiarowych jest tak mały, że staje się porównywalny z czasem relaksacji zderzeniowej, jak również wówczas, gdy rozmiary obszaru, w którym odbywa się transport nośników, stają się porównywalne z ich średnią drogą swobodną. Przypadki te wymagają sformułowania równań uwzględniających mikroskopowy charakter zjawisk transportu. Zagadnieniami tymi nie będziemy się zajmowali w tej książce.

Ostatni punkt tego rozdziału jest poświęcony przykładom rozwiązywania równań transportu. Będą tu rozważone 2 zagadnienia, mające istotne znaczenie w teorii elementów półprzewodnikowych.

12.1 RÓWNANIE CIĄGŁOŚCI

Jednym z ważnych wniosków wypływających z równań Maxwella jest *równanie ciągłości* przedstawione w postaci (A/8) w dodatku A. Równanie to jest wyrazem podstawowego prawa przyrody – *zasady zachowania ładunku*[1]. Orzeka ono,



12/1
Jednowymiarowy przepływ prądu w półprzewodniku

że ładunek elektryczny dowolnego układu pozostaje niezmienny w czasie, jeżeli układ ten jest elektrycznie odosobniony. Ładunek ten może więc zmieniać się tylko wtedy, gdy przez powierzchnię otaczającą układ przepływa prąd elektryczny. Prawo to dla nieskończenie małego obszaru przybiera postać

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot \boldsymbol{J} = 0 \qquad (12/1)$$

gdzie: ϱ – gęstość ładunku przestrzennego; $\boldsymbol{J}$ – gęstość prądu elektrycznego.

Dla przepływu jednowymiarowego można równanie ciągłości wyprowadzić na podstawie rys. 12/1. Weźmy pod uwagę zaznaczony na rysunku obszar ogra-

[1] Prawo to odnosi się nie tylko do ładunków elekuycznych, ale znajduje zastosowanie również w innych przypadkach, np. w mechanice płynów równanie ciągłości jest wyrazem zasady zachowania masy poruszającego się płynu.

niczony płaszczyznami x oraz $x+dx$, i oznaczmy symbolem A pole jego przekroju poprzecznego. Prąd wpływający do tego obszaru jest równy $AJ(x)$, a prąd wypływający z niego $-AJ(x+dx)$. Z kolei ładunek zawarty w tym obszarze wynosi $A\varrho(x)\,dx$. Ze względu na to, że ładunki elektryczne nie mogą ani ginąć, ani też powstawać z niczego, szybkość zmian ładunku zawartego w danym obszarze musi być równa różnicy prądów wpływającego i wypływającego. Stwierdzenie to można wyrazić wzorem

$$A\frac{\partial\varrho(x)}{\partial t}\,dx = A\,[J(x)-J(x+dx)] \tag{12/2}$$

z którego wynika, że

$$\boxed{\frac{\partial\varrho}{\partial t} = -\frac{\partial J}{\partial x}} \tag{12/3}$$

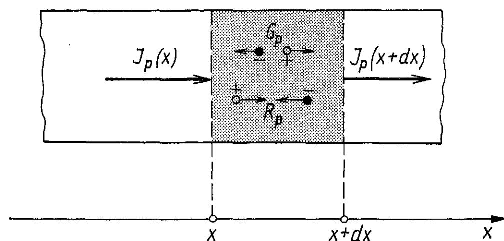
Otrzymany wzór jest równoważny wyrażeniu (12/1) przy założeniu, że przepływy ładunków zachodzą w kierunku równoległym do osi x i że zarówno ϱ, jak i J zależą tylko od tej jednej współrzędnej przestrzennej i, oczywiście, od czasu.

W rozdziale tym będziemy rozpatrywali tylko przepływy jednowymiarowe. Czytelnik może jednak bez trudu rozszerzyć te rozważania na przepływy wielowymiarowe.

Równanie ciągłości odnosi się do całkowitego ładunku elektrycznego znajdującego się w rozważanym obszarze. Nie można go natomiast w tej formie stosować do jednego rodzaju nośników w półprzewodniku. Elektrony i dziury nie podlegają bowiem zasadzie zachowania, mogą bowiem ulegać anihilacji wskutek rekombinacji, lub też powstawać „z niczego" w wyniku generacji.

Chcąc zastosować równanie ciągłości do jednego rodzaju nośników, np. do dziur, należy w nim uwzględnić wszystkie mechanizmy mające wpływ na ilość tych nośników – zob. rys. 12/2. Liczba dziur w rozpatrywanym obszarze może wzrosnąć w wyniku przewagi dopływu dziur nad ich odpływem oraz w wyniku przewagi generacji nad rekombinacją. Zamiast równania (12/2) musimy więc napisać

$$Aq\frac{\partial p}{\partial t}\,dx = A\,[J_p(x)-J_p(x+dx)]+Aq\,(G_p-R_p)\,dx \tag{12/4}$$



12/2 Czynniki wpływające na koncentrację dziur przy przepływie jednowymiarowym

Stąd otrzymuje się równanie ciągłości dla dziur

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} + G_p - R_p \tag{12/5}$$

przy czym G_p i R_p oznaczają odpowiednio szybkości generacji i rekombinacji dziur.

Szybkości te różnią się na ogół od analogicznych wielkości G_n i R_n dla elektronów, a przyczyną tego jest omawiane w p. 9.3.3 pułapkowanie nośników. Równanie ciągłości dla elektronów ma zatem postać

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} + G_n - R_n \tag{12/6}$$

Zjawisko pułapkowania nośników jest trudne do uwzględnienia, a ponadto odgrywa ono zazwyczaj mało istotną rolę. Dlatego też w dalszych rozważaniach będziemy przyjmowali, że stopień obsadzenia pułapek w półprzewodniku ulega tak niewielkim zmianom, że praktycznie rzecz biorąc $G_n = G_p$ oraz $R_n = R_p$. Przy takim założeniu można różnice $G_p - R_p = G_n - R_n$ wyrazić za pomocą różnicy $g - U_{rg}$, gdzie: g – szybkość generacji nadmiarowych par elektron-dziura (wywołanej np. promieniowaniem elektromagnetycznym), a U_{rg} – zdefiniowana w rozdz. 9 wypadkowa szybkość rekombinacji nośników. Równania ciągłości dla elektronów i dziur można wówczas przedstawić w postaci:

$$\frac{\partial n}{\partial t} = g + \frac{1}{q}\frac{\partial J_n}{\partial x} - U_{rg} \tag{12/7}$$

$$\frac{\partial p}{\partial t} = g - \frac{1}{q}\frac{\partial J_p}{\partial x} - U_{rg} \tag{12/8}$$

Równania te są ogólnie słuszne, odnoszą się więc zarówno do półprzewodników jednorodnych, jak i niejednorodnych o dowolnym rodzaju niejednorodności (a więc również np. do złącza *p-n*). Są to równania różniczkowe cząstkowe, ponieważ występują w nich pochodne cząstkowe względem współrzędnej x i względem czasu.

W celu umożliwienia rozwiązania równań ciągłości należy uzależnić wielkości występujące po prawej stronie równości (12/7) i (12/8) od koncentracji elektronów i dziur. Korzystając z wyprowadzonych w rozdz. 10 wyrażeń określających gęstości prądów J_n i J_p (wzory (10/64) i (10/65)), otrzymamy:

$$\frac{\partial n}{\partial t} = g + \mu_n \frac{\partial}{\partial x}(nE) + D_n \frac{\partial^2 n}{\partial x^2} - U_{rg} \tag{12/9}$$

$$\frac{\partial p}{\partial t} = g - \mu_p \frac{\partial}{\partial x}(pE) + D_p \frac{\partial^2 p}{\partial x^2} - U_{rg} \tag{12/10}$$

Jako wypadkową szybkość rekombinacji U_{rg} należy tu podstawić odpowiednie wyrażenie z rozdz. 9. Jest to zwykle dość skomplikowana funkcja koncentracji nośników, która tylko w pewnych warunkach daje się uprościć do postaci typu (9/16).

Przy różniczkowaniu wyrażeń (10/64) i (10/65) założyliśmy milcząco, że parametry materiałowe μ_n, μ_p, D_n i D_p są niezależne od współrzędnej x. Ściśle biorąc, założenie to jest słuszne tylko dla półprzewodnika jednorodnego w warunkach izotermicznych. W półprzewodniku, w którym koncentracje domieszek zmieniają się wzdłuż osi x, należałoby uwzględnić zależność ruchliwości (i związanych z nimi współczynników dyfuzji) od koncentracji domieszek. W przypadku natomiast nierównomiernego rozkładu temperatury w półprzewodniku należałoby nie tylko uwzględnić zależność współczynników μ i D od temperatury, lecz również uzupełnić wzory (10/64) i (10/65) odpowiednimi składnikami proporcjonalnymi do gradientu temperatury.

W równaniach (12/9) i (12/10) występują 3 nieznane funkcje $n(x, t)$, $p(x, t)$ i $E(x, t)$. Do rozwiązania dowolnego zagadnienia dotyczącego przepływu nośników nadmiarowych w półprzewodniku jest więc potrzebne jeszcze jedno niezależne równanie opisujące związek zachodzący między tymi trzema wielkościami. Jest nim poznane w rozdz. 11 *równanie Poissona*, które tu dla wygody powtórzymy. Podstawiając (11/8) do (11/7b) i uwzględniając to, że natężenie pola jest nie tylko funkcją współrzędnej x, lecz zależy również od czasu, otrzymujemy

$$\boxed{\frac{\partial E}{\partial x} = \frac{q}{\varepsilon}(N_d^+ - N_a^- + p - n)} \qquad (12/11)$$

Układ równań (12/9) ... (12/11) umożliwia w zasadzie wyznaczenie szukanych wielkości n, p i E w postaci funkcji współrzędnej x i czasu. Rozwiązania otrzymane dla określonego przypadku muszą przy tym spełniać odpowiednie warunki brzegowe i początkowe. Mimo wprowadzonych uproszczeń, analityczne rozwiązanie tych równań jest niemożliwe nawet dla najprostszych niebanalnych przypadków. Konieczne jest zatem albo zastosowanie dalszych uproszczeń, dzięki którym znalezienie rozwiązania analitycznego dla prostych sytuacji staje się możliwe, albo też posługiwanie się metodami numerycznymi. Wówczas jednak tracimy na ogólności rozwiązania. Wyniki uzyskane metodami numerycznymi odnoszą się bowiem do konkretnych przypadków i nie mogą być przedstawione w postaci ogólnych zależności między interesującymi nas wielkościami.

W zakończeniu tego punktu warto jeszcze raz zaznaczyć, że równania (12/9)–(12/11) opierają się na następujących założeniach upraszczających:

1) jednowymiarowy przepływ nośników;

2) pomijalnie mały wpływ pułapkowania na szybkość generacji i rekombinacji nośników;

3) równomierny rozkład temperatury;

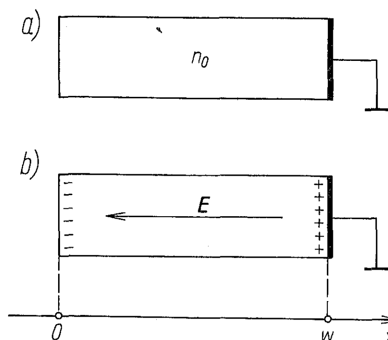
4) niezależność parametrów materiałowych półprzewodnika od współrzędnej x.

W dalszych rozważaniach zawartych w tym rozdziale założymy ponadto, że rozpatrywany półprzewodnik jest jednorodny. Jest to, jak już zaznaczono, warunek wystarczający do tego, by można było założenie 4) uznać za spełnione. Założymy również, że wszystkie domieszki donorowe i akceptorowe są zjonizowane. W tych warunkach równanie Poissona przyjmuje postać (11/10b).

12.2 RELAKSACJA DIELEKTRYCZNA

12.2.1 Zakłócenie rozkładu koncentracji nośników większościowych

Rozpatrzmy jednorodny półprzewodnik, w którym pod działaniem zewnętrznego bodźca nastąpiło lokalne zakłócenie rozkładu koncentracji nośników większościowych. Przyjmijmy przykładowo, że jest to półprzewodnik typu n w postaci prostopadłościennej próbki zakończonej idealnym kontaktem rekombinacyjnym na potencjale ziemi. Półprzewodnik ten znajduje się początkowo w stanie równowagi termodynamicznej – rys. 12/3a.



12/3
Zakłócenie rozkładu koncentracji nośników większościowych: a) próbka półprzewodnika typu *n*; b) stan zakłócony

Założmy, że w chwili $t = 0$ do półprzewodnika została z zewnątrz wprowadzona pewna liczba dodatkowych elektronów, tak jak to pokazano na rys. 12/3b. To zakłócenie rozkładu ładunków elektrycznych staje się źródłem fali elektromagnetycznej, która rozprzestrzenia się z prędkością światła w półprzewodniku i w otaczającej przestrzeni. W wyniku tego powstaje w półprzewodniku pole elektryczne, skierowane od ładunków dodatnich, indukowanych na powierzchni kontaktu, do ładunków ujemnych, utworzonych przez elektrony wprowadzone do warstwy przypowierzchniowej półprzewodnika. Pole to oddziałuje na wszystkie nośniki w półprzewodniku i wywołuje przepływ prądu unoszenia. Stan ten trwa tak długo, jak długo istnieją nośniki nadmiarowe i wytworzone przez nie pole elektryczne. Jak zobaczymy na końcu tego wywodu, czas występowania zakłócenia jest zazwyczaj niezwykle krótki. Z tego względu w rozważaniach naszych możemy nie brać pod uwagę procesów dyfuzji i rekombinacji nośników.

Równanie prądu i równanie ciągłości mają zatem postać:

$$J = \sigma E \tag{12/12}$$

$$\frac{\partial \Delta n}{\partial t} = \frac{1}{q} \frac{\partial J_n}{\partial x} \tag{12/13}$$

przy czym można przyjąć, że $J = J_n$, ponieważ w dostatecznie silnie domieszkowanym półprzewodniku typu n składnik dziurowy prądu unoszenia jest pomijalnie mały.

Korzystając następnie z równania Poissona

$$\frac{\partial E}{\partial x} = -\frac{q}{\varepsilon} \Delta n \tag{12/14}$$

otrzymuje się równanie opisujące zmiany w czasie nadmiarowej koncentracji elektronów Δn w dowolnym punkcie x

$$\frac{\partial \Delta n}{\partial t} = -\frac{\sigma}{\varepsilon} \Delta n \tag{12/15}$$

Rozwiązanie tego równania ma postać

$$\Delta n (t) = \Delta n (0) \exp\left(-\frac{t}{\tau_d}\right) \tag{12/16}$$

Nadmiarowa koncentracja nośników większościowych zanika więc w czasie wykładniczo od wartości początkowej $\Delta n (0)$ ze stałą czasową

$$\boxed{\tau_d = \frac{\varepsilon}{\sigma}} \tag{12/17}$$

którą nazywamy *czasem relaksacji dielektrycznej.*

Podstawiając do wzoru (12/17) odpowiednie wartości stałych materiałowych dla krzemu o koncentracji domieszek $N_d = 10^{16}$ cm^{-3} ($\varepsilon = 10^{-12}$ F/cm, $\sigma = 1\ (\Omega \cdot \mathrm{cm})^{-1}$), otrzymujemy $\tau_d = 10^{-12}$ s. Jest to istotnie czas bardzo krótki, a zatem przyjęte założenia o nieuwzględnieniu procesów dyfuzji i rekombinacji nośników można uznać za w pełni uzasadnione.

Z rozważań tych wynika, że każde zakłócenie rozkładu koncentracji nośników większościowych zostaje niemal natychmiast zlikwidowane. Dzieje się tak, ponieważ pole elektryczne powstające wskutek tego zakłócenia oddziałuje na wszystkie nośniki większościowe i powoduje takie ich przemieszczenie, że rozkład koncentracji tych nośników powraca do rozkładu równowagowego. Nadmiar nośników zostaje usunięty przez kontakt, którym jest zakończona próbka półprzewodnika.

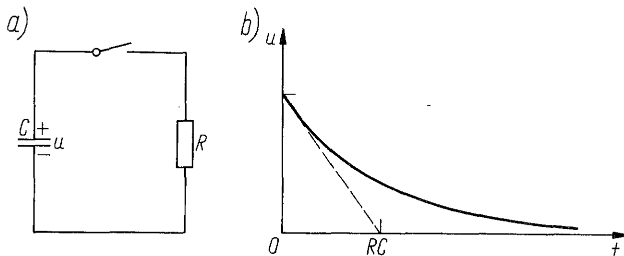
Proces relaksacji dielektrycznej zachodzi tym szybciej, im większa jest koncentracja i ruchliwość nośników większościowych, a więc im jest większa przewodność półprzewodnika.

Problem 12.1. Załóżmy, że próbka półprzewodnika przedstawiona na rys. 12/3 jest prostopadłościanem o polu podstawy (prostopadłej do osi x) równym 1 cm^2, i że jest ona wykonana z krzemu o koncentracji domieszek $N_d = 10^{16}$ cm^{-3}. Przez podstawę wprowadzono

do półprzewodnika $6 \cdot 10^8$ elektronów. Wyznacz wartość natężenia pola elektrycznego, które powstanie w wyniku tego zakłócenia oraz wartość prądu, który popłynie w chwili wytworzenia się pola elektrycznego. Jak długo płynąłby ten prąd, gdyby jego natężenie nie ulegało zmianie przez czas trwania zakłócenia? Przedyskutuj otrzymane wyniki.

Proces relaksacji dielektrycznej w półprzewodniku zachodzi w podobny sposób, jak rozładowanie kondensatora w układzie przedstawionym na rys. 12/4. To ostatnie odbywa się, jak wiadomo, ze stałą czasową

$$\tau = RC \qquad (12/18)$$



12/4 Proces rozładowania kondensatora: a) schemat układu elektrycznego; b) zależność napięcia od czasu

Stosując ten wzór do rozpatrywanej próbki półprzewodnika, należy podstawić

$$R = \frac{1}{\sigma}\frac{w}{A} \qquad (12/19)$$

$$C = \varepsilon\frac{A}{w} \qquad (12/20)$$

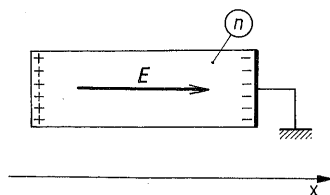
gdzie: A – pole podstawy; w – długość próbki.

Po podstawieniu tych wyrażeń do wzoru (12/18) otrzymuje się wynik identyczny z uzyskanym uprzednio na podstawie rozwiązania równań transportu nośników.

12.2.2 Zakłócenie rozkładu koncentracji nośników mniejszościowych

Zajmiemy się teraz zagadnieniem analogicznym do rozpatrywanego w poprzednim punkcie z tą jednak różnicą, że zamiast nośników większościowych zostaną wprowadzone do półprzewodnika nadmiarowe nośniki mniejszościowe, tzn. dziury. W półprzewodniku powstanie wówczas pole elektryczne o kierunku pokazanym na rys. 12/5, a pod jego działaniem nastąpi przemieszczenie się nośników, głównie nośników większościowych, tzn. elektronów.

Nośniki te przesuną się w kierunku przeciwnym do kierunku wektora $\boldsymbol{E}$, a więc, w kierunku podstawy, przez którą zostały wprowadzone nadmiarowe dziury. Starają się one skompensować dodatni ładunek, będący źródłem pola



12/5 Zakłócenie rozkładu koncentracji nośników mniejszościowych w półprzewodniku typu *n*

elektrycznego. Pole to przestałoby istnieć, gdyby w każdym punkcie półprzewodnika był spełniony warunek neutralności

$$\Delta n(x) = \Delta p(x) \qquad (12/21)$$

W rzeczywistości warunek ten jest spełniony tylko w przybliżeniu, ponieważ niewielka część początkowego pola musi pozostać, aby utrzymać rozkład koncentracji nośników większościowych odbiegający od rozkładu w stanie równowagi termodynamicznej. Podobnie jak w rozdziale 11 mamy tu więc do czynienia z warunkiem *quasi-neutralności elektrycznej.* Koncentracje nadmiarowych elektronów i nadmiarowych dziur różnią się przy tym bardzo nieznacznie, ponieważ niewielka ilość nieskompensowanych ładunków wystarcza do wytworzenia pola elektrycznego o natężeniu potrzebnym do utrzymania nierównowagowego rozkładu nośników większościowych[1].

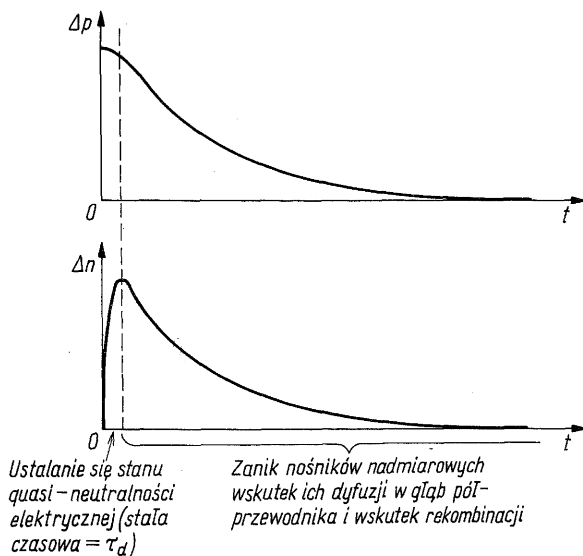
Ustalenie się warunków quasi-neutralności elektrycznej odbywa się wg wykładniczej funkcji czasu ze stałą czasową równą τ_d. Dla półprzewodników takich jak krzem, dla których czas relaksacji dielektrycznej jest rzędu pikosekundy, można więc przyjmować, że stan quasi-neutralności elektrycznej zostaje osiągnięty niemal natychmiast.

Stan ten różni się jednakże od warunków równowagi termodynamicznej, w półprzewodniku występują bowiem nośniki nadmiarowe obu rodzajów, przy czym rozkład ich koncentracji jest niemal identyczny. Taki sam rozkład można by otrzymać, wywołując generację par elektron-dziura, np. przez oświetlenie półprzewodnika krótkotrwałym impulsem promieniowania elektromagnetycznego. W tym sensie wprowadzenie do półprzewodnika nadmiarowego nośników mniejszościowych jest więc równoważne generacji nadmiarowych par elektron-dziura.

Po osiągnięciu stanu quasi-neutralności elektrycznej rozpoczyna się druga faza powrotu półprzewodnika do stanu równowagi termodynamicznej[2]. Wskutek nierównomiernego rozkładu koncentracji nośników nadmiarowych następuje ich dyfuzja połączona z unoszeniem pod działaniem sił pola elektrycznego. Niezależnie od tego przebiegają procesy rekombinacji, które przy małym poziomie zakłócenia zachodzą ze stałą czasową równą czasowi życia nośników mniej-

[1] Jak zobaczymy w p. 12.5.2 w przypadku krzemu koncentracje nadmiarowe mogą się różnić np. o jedną milionową.

[2] Ściśle biorąc, omawiane tu procesy rozpoczynają się równocześnie. Jednakże, z uwagi na bardzo małą wartość czasu relaksacji dielektrycznej, można umownie przyjmować, że proces powrotu do stanu równowagi odbywa się w dwóch etapach, różniących się wyraźnie dominacją jednych zjawisk nad innymi.



12/6
Zmiany rozkładu koncentracji nośników nadmiarowych w funkcji czasu w próbce z rys. 12/5

szościowych (w rozpatrywanym przypadku – τ_p). Czas ten jest dla krzemu o kilka rzędów wielkości większy niż czas relaksacji dielektrycznej. Procesy te doprowadzają asymptotycznie do zaniku nośników nadmiarowych obu rodzajów, a więc do ustalenia się równowagi termodynamicznej – zob. rys. 12/6.

Powyższe rozważania pozwalają zrozumieć, dlaczego w półprzewodniku jest możliwe wywołanie dużych zmian rozkładu koncentracji nośników. Gdyby w półprzewodniku występowały tylko nośniki jednego rodzaju, jak to np. zachodzi w metalach, wówczas wprowadzenie dodatkowych nośników spowodowałoby powstanie nieskompensowanego ładunku elektrycznego i związanego z tym pola elektrycznego. W rezultacie, wprowadzony ładunek zanikłby bardzo szybko – proces zaniku zachodziłby bowiem ze stałą czasową równą czasowi relaksacji dielektrycznej.

W półprzewodniku występują jednakże 2 rodzaje nośników i, jak już stwierdziliśmy, wprowadzenie nadmiarowych nośników mniejszościowych niemal nie wywołuje zakłócenia neutralności elektrycznej. Z tego względu powrót do stanu równowagi termodynamicznej zachodzi tu wskutek rekombinacji par elektron-dziura. Jest to zazwyczaj proces stosunkowo powolny, a więc stan zakłócenia rozkładu koncentracji nośników utrzymuje się w półprzewodniku dość długo. W tych warunkach bardzo ważnym zjawiskiem występującym w półprzewodnikach staje się dyfuzja nośników, która natomiast w metalach nie odgrywa istotnej roli.

Rozważania te odnoszą się do półprzewodników, dla których czas życia nośników nadmiarowych jest o kilka rzędów wielkości dłuższy niż czas relaksacji dielektrycznej. Tylko w takich półprzewodnikach (zwanych niekiedy półprzewodnikami *rekombinacyjnymi*) zjawiska związane z obecnością nośników nadmiarowych zachodzą w warunkach prawie zachowanej neutralności elektrycznej.

Istnieją jednakże półprzewodniki, do których należy np. półizolacyjny GaAs oraz półprzewodniki amorficzne, mające czasy relaksacji dielektrycznej znacznie dłuższe niż czasy życia nośników nadmiarowych. Dla takich półprzewodników, zwanych półprzewodnikami *relaksacyjnymi*, zjawiska opisane w tym punkcie przebiegają w sposób odmienny. Różnica polega m.in. na tym, że w półprzewodnikach relaksacyjnych warunek quasi-neutralności elektrycznej nie jest spełniony. Zagadnienia te wykraczają jednakże poza ramy tej książki, w której będziemy się zajmowali jedynie półprzewodnikami rekombinacyjnymi[1].

12.3 AMBIPOLARNE RÓWNANIE TRANSPORTU

Dla półprzewodnika jednorodnego równania (12/9)–(12/11) upraszczają się do postaci

$$\frac{\partial \Delta n}{\partial t} = g + n\mu_n \frac{\partial E}{\partial x} + \mu_n E \frac{\partial \Delta n}{\partial x} + D_n \frac{\partial^2 \Delta n}{\partial x^2} - U_{rg} \tag{12/22}$$

$$\frac{\partial \Delta p}{\partial t} = g - p\mu_p \frac{\partial E}{\partial x} - \mu_p E \frac{\partial \Delta p}{\partial x} + D_p \frac{\partial^2 \Delta p}{\partial x^2} - U_{rg} \tag{12/23}$$

$$\frac{\partial E}{\partial x} = \frac{q}{\varepsilon}(\Delta p - \Delta n) \tag{12/24}$$

Uproszczenia te wynikają z podstawienia do równań (12/9)–(12/11) wyrażeń (9/1) i (9/2), oraz uwzględnienia faktu, że w półprzewodniku jednorodnym koncentracje równowagowe n_0 i p_0 nie zależą ani od czasu, ani od współrzędnej x. Wzór (12/24) jest identyczny ze wzorem (11/10b) – z tą uwagą, że występuje w nim pochodna cząstkowa $\partial E/\partial x$, ponieważ w rozdziale tym zajmujemy się przebiegami zależnymi nie tylko od współrzędnej x, lecz również od czasu.

Zgodnie z wynikami rozważań poprzedniego punktu można przyjąć, że po upływie czasu wynoszącego kilka τ_d w każdym punkcie półprzewodnika i w każdej chwili jest spełniony warunek quasi-neutralności elektrycznej, tzn.

$$\Delta n\,(x, t) \approx \Delta p\,(x, t) \tag{12/25}$$

Nie oznacza to jednak, że prawą stronę równania (12/24) można przyrównać do zera. Różnica $(\Delta p - \Delta n)$ jest bowiem mała w stosunku do koncentracji nośników nadmiarowych, nie może jednak być pominięta. Podobnie jak w przypadku rozpatrywanym w p. 11.3, różnica ta jest źródłem wewnętrznego pola elektrycznego w półprzewodniku, a jak dalej zobaczymy, pole to odgrywa w wielu przypadkach istotną rolę.

Ze względu na trudność wyznaczenia wielkości $(\Delta p - \Delta n)$, spróbujemy w inny sposób pozbyć się w równaniach (12/22) i (12/23) składnika zawierającego pochodną $\partial E/\partial x$. Pomnożymy w tym celu pierwsze z tych równań przez $p\mu_p$,

---

[1] Zjawiska transportu nośników w półprzewodnikach relaksacyjnych są omawiane np. w książce: J. Marciak-Kozłowska: *Niestacjonarne zjawiska transportu w półprzewodnikach*. PWN, Warszawa 1982.

drugie zaś – przez $n\mu_n$ i następnie dodamy je stronami. Uwzględniając (12/25), otrzymamy wówczas

$$\frac{\partial \Delta p}{\partial t} = g - \frac{\mu_n \mu_p (n-p)}{n\mu_n + p\mu_p} E \frac{\partial \Delta p}{\partial x} + \frac{n\mu_n D_p + p\mu_p D_n}{n\mu_n + p\,\mu_p} \frac{\partial^2 \Delta p}{\partial x^2} - U_{rg} \quad (12/26)$$

Otrzymane równanie można przedstawić w postaci

$$\boxed{\frac{\partial \Delta p}{\partial t} = g - \mu^* E \frac{\partial \Delta p}{\partial x} + D^* \frac{\partial^2 \Delta p}{\partial x^2} - U_{rg}} \quad (12/27)$$

Nosi ono nazwę *ambipolarnego równania transportu*, ponieważ opisuje zachowanie się obu rodzajów nośników nadmiarowych[1]. Formalnie równanie to dotyczy dziur, jednakże ze względu na warunek (12/25) opisuje ono także zachowanie się nadmiarowych elektronów. Oddziaływania występujące między tymi nośnikami są zawarte we wprowadzonych tu nowych parametrach μ^* i D^*, zdefiniowanych w następujący sposób:

$$\mu^* = \frac{n-p}{\dfrac{n}{\mu_p} + \dfrac{p}{\mu_n}} \quad (12/28)$$

$$D^* = \frac{n+p}{\dfrac{n}{D_p} + \dfrac{p}{D_n}} \quad (12/29)$$

Pierwszy z nich nosi nazwę *ruchliwości grupowej*, drugi zaś – *ambipolarnego współczynnika dyfuzji*. Nazwa „ruchliwość grupowa" jest usprawiedliwiona tym, że parametr μ^* jest związany z przemieszczaniem się impulsu zakłócenia koncentracji elektronów i dziur pod działaniem pola elektrycznego. Nośniki nadmiarowe, obdarzone ładunkami przeciwnego znaku, są ze sobą tak silnie związane siłami elektrycznymi, że w polu elektrycznym obserwujemy pozorne przesuwanie się nierozdzielnej grupy tych nośników. Zachowanie się tej grupy jest opisane za pomocą ambipolarnego równania transportu. Same nośniki poruszają się natomiast w inny sposób. Pod działaniem pola elektrycznego dziury poruszają się bowiem w przeciwnym kierunku niż elektrony. Zagadnienie to będzie dokładniej omówione w p. 12.4.

Jak wynika ze wzorów (12/28) i (12/29) parametry μ^* i D^* zależą od odpowiednich parametrów μ_n i μ_p oraz D_n i D_p, a także od koncentracji elektronów i dziur. Ruchliwość grupowa zależy przy tym od różnicy $n-p$, natomiast ambipolarny współczynnik dyfuzji od sumy $n+p$. Wynika to stąd, że pod wpływem pola elektrycznego elektrony mają tendencję do poruszania się w kierunku przeciwnym niż dziury, a zatem obecność dziur przeciwdziała unoszeniu elektronów,

[1] Nazwa „ambipolarny" oznacza dwubiegowy – od łacińskiego słowa „ambo" (obaj) i greckiego „polos" (biegun).

dyfuzja natomiast elektronów i dziur zachodzi w tym samym kierunku, a więc przeciwdziałanie takie nie występuje.

Dla dostatecznie silnie domieszkowanych półprzewodników przy małym poziomie zakłócenia wyrażenia (12/28) i (12/29) upraszczają się znacznie. Dla półprzewodnika typu n mamy wówczas bowiem $n \gg p$, a zatem

$$\mu^* = \mu_p \tag{12/30}$$

$$D^* = D_p \tag{12/31}$$

Jednocześnie, w podanych tutaj warunkach wypadkowa szybkość rekombinacji wyraża się wzorem

$$U_{rg} = \frac{\Delta p}{\tau_p} \tag{12/32}$$

Ambipolarne równanie transportu przyjmuje zatem postać

$$\boxed{\frac{\partial \Delta p}{\partial t} = g - \mu_p E \frac{\partial \Delta p}{\partial x} + D_p \frac{\partial^2 \Delta p}{\partial x^2} - \frac{\Delta p}{\tau_p}} \tag{12/33}$$

Zachowanie się nośników nadmiarowych zależy więc od ruchliwości, współczynnika dyfuzji i czasu życia nośników *mniejszościowych*. Własności nośników większościowych są tu ukryte w wielkości E reprezentującej natężenie pola elektrycznego. Pole to jest sumą pola wytwarzanego przez ładunki zewnętrzne oraz pola, którego źródłem jest ładunek przestrzenny w półprzewodniku. Wyznaczenie tego ostatniego jest jednym z trudniejszych problemów w rozwiązywaniu zagadnień transportu nośników w półprzewodniku. W wielu przypadkach można jednakże z niewielkim błędem zaniedbać pole wytwarzane przez ładunek przestrzenny. Dla małego poziomu zakłócenia jest to powszechnie stosowane założenie upraszczające.

Dla półprzewodnika typu p, przy małym poziomie zakłócenia, zachodzi nierówność $p \gg n$, co sprawia, że

$$\mu^* = -\mu_n \tag{12/34}$$

$$D^* = D_n \tag{12/35}$$

Jednocześnie

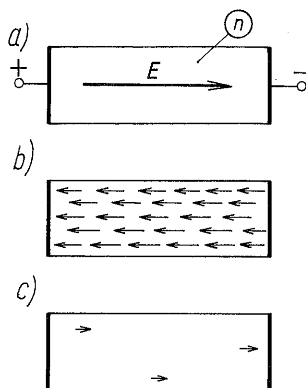
$$U_{rg} = \frac{\Delta n}{\tau_n} \tag{12/36}$$

a zatem ambipolarne równanie transportu przyjmuje postać

$$\boxed{\frac{\partial \Delta n}{\partial t} = g + \mu_n E \frac{\partial \Delta n}{\partial x} + D_n \frac{\partial^2 \Delta n}{\partial x^2} - \frac{\Delta n}{\tau_n}} \tag{12/37}$$

12.4 DYSKUSJA AMBIPOLARNEGO RÓWNANIA TRANSPORTU

W celu lepszego zrozumienia ambipolarnego równania transportu, rozpatrzmy zachowanie się nośników nadmiarowych w prostopadłościennej próbce półprzewodnika zakończonej kontaktami, do których zostało doprowadzone stałe napięcie – zob. rys. 12/7a. Załóżmy, że mamy do czynienia z półprzewodnikiem typu n i przeanalizujemy najpierw przepływ nośników w warunkach, w których nie ma nośników nadmiarowych.



12/7
Przepływ nośników w półprzewodniku typu *n*: a) próbka półprzewodnika; b) przepływ elektronów; c) przepływ dziur

W półprzewodniku występuje wówczas jednorodne pole elektryczne, którego natężenie oznaczymy symbolem E_0. Pod działaniem tego pola elektrony poruszają się w kierunku kontaktu o potencjale dodatnim, dziury zaś – w kierunku przeciwnym, jak to w sposób poglądowy pokazano na rys. 12/7b i c.

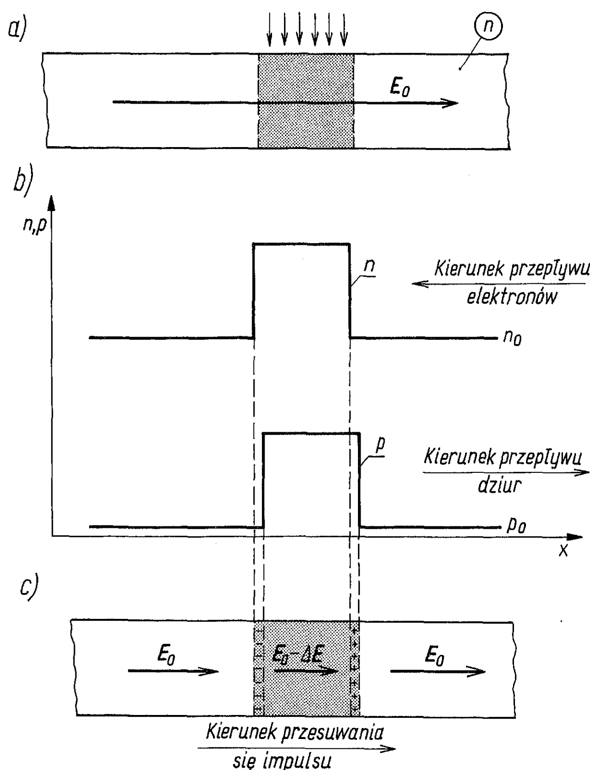
Załóżmy teraz, że w próbce półprzewodnika zostały wytworzone nośniki nadmiarowe, np. w wyniku krótkotrwałego oświetlenia środkowego odcinka próbki promieniowaniem elektromagnetycznym o dostatecznie dużej energii fotonów – zob. rys. 12/8a. W celu ułatwienia analizy, przyjmijmy, że rozkład koncentracji nośników nadmiarowych ma kształt impulsu prostokątnego, przy czym długość tego impulsu jest mała w stosunku do długości próbki półprzewodnika.

W wyniku działania pola elektrycznego nastąpi niewielkie rozsunięcie impulsów nadmiarowych elektronów i nadmiarowych dziur – zob. rys. 12/8b. Pole elektryczne „odciąga" bowiem dziury w kierunku zgodnym z kierunkiem wektora E_0, elektrony natomiast – w kierunku przeciwnym. To rozsunięcie impulsów nośników sprawia, że w półprzewodniku powstają dwie warstwy nieskompensowanych ładunków (rys. 12/8c), które stają się źródłem dodatkowego pola elektrycznego, skierowanego przeciwnie niż pole wytworzone pod wpływem doprowadzonego napięcia zewnętrznego.

W obszarze, w którym występuje impuls nośników nadmiarowych, pole elektryczne jest więc słabsze; jego natężenie oznaczymy symbolem $E_0-\Delta E$. Gęstości prądu elektronowego i dziurowego wynoszą tu odpowiednio

$$J_n = q\,(n_0+\Delta n)\,\mu_n(E_0-\Delta E) \tag{12/38}$$

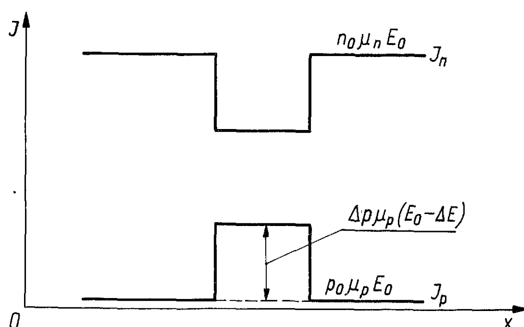
$$J_p = q\,(p_0+\Delta p)\,\mu_p(E_0-\Delta E) \tag{12/39}$$



12/8
Przepływ impulsu nośników nadmiarowych w półprzewodniku typu n: a) impuls nośników nadmiarowych; b) rozkład koncentracji nośników; c) rozkład pola elektrycznego

przy czym w środkowej części impulsu $\Delta n = \Delta p$. Przebieg zależności J_n i J_p od współrzędnej x pokazano w pewnym uproszczeniu na rys. 12/9. Jak widać na rysunku, gęstość prądu wywołanego unoszeniem nośników większościowych (w rozważanym przypadku – elektronów) jest w obszarze impulsu mniejsza niż poza tym obszarem. Jest to spowodowane tym, że wobec dużej wartości n_0, względna zmiana koncentracji tych nośników jest mniejsza niż względna zmiana natężenia pola elektrycznego.

Gęstość prądu wywołanego unoszeniem nośników mniejszościowych (dziur) jest natomiast w obszarze impulsu większa niż poza tym obszarem. W tym przy-



12/9
Rozkład gęstości prądu elektronowego i dziurowego dla przypadku przedstawionego na rys. 12/8

padku bowiem względna zmiana koncentracji nośników przewyższa (często nawet bardzo znacznie) względną zmianę natężenia pola elektrycznego.

Narastanie dodatkowego pola elektrycznego, którego źródłem jest ładunek przestrzenny w półprzewodniku, zachodzi tak długo, jak długo suma wyrażeń (12/38) i (12/39) nie zrówna się z gęstością prądu unoszenia poza obszarem impulsu nośników nadmiarowych. Wynika to z faktu, że gęstość prądu unoszenia w próbce musi być niezależna od współrzędnej x[1]. Stwierdzenie to można wyrazić równością

$$J_n(x)+J_p(x) = \text{const} \qquad (12/40)$$

Problem 12.2. Opierając się na wzorze (12/40), wyprowadź wzór na wielkość ΔE, o jaką zmniejsza się natężenie pola elektrycznego w obszarze impulsu nośników nadmiarowych. Sporządź wykresy zależności $E(x)$, $J_n(x)$ i $J_p(x)$ dla następujących przypadków: a) półprzewodnik samoistny, $\Delta n = 20\, n_i$; *b)* silnie domieszkowany półprzewodnik typu n, $\Delta n = 0{,}07\, n_0$. W obu przypadkach przyjmij, że $\mu_n = 3\, \mu_p$. Przedyskutuj otrzymane wyniki.

Ze wzoru otrzymanego w problemie 12.2 wynika, że dla dostatecznie silnie domieszkowanego półprzewodnika typu n przy małym poziomie zakłócenia stosunek $\Delta E/E_0$ jest tego samego rzędu wielkości co stosunek $\Delta n/n_0$. Pole elektryczne w obszarze impulsu nośników nadmiarowych różni się więc tylko nieznacznie od pola, które występowało w półprzewodniku w warunkach, gdy nie było w nim nośników nadmiarowych.

Nadmiarowe dziury są unoszone przez pole z prędkością równą $\mu_p(E_0-\Delta E)$, a zatem impuls nadmiarowych dziur przesuwa się z prędkością tylko nieznacznie odbiegającą od wartości $\mu_p E_0$. Tym tłumaczy się wyprowadzony w poprzednim punkcie wniosek, że dla dostatecznie silnie domieszkowanego półprzewodnika, przy małym poziomie zakłócenia, ruchliwość grupowa jest, co do wartości bezwzględnej, równa ruchliwości nośników *mniejszościowych.*

Pozostaje jeszcze do wyjaśnienia problem, dlaczego impuls nośników *większościowych* przesuwa ię wraz z nadmiarowymi nośnikami mniejszościowymi, a więc w kierunku przeciwnym do kierunku przepływu tych pierwszych. Jest to konsekwencja warunku (12/40), z którego wynika określony rozkład pola ładunku przestrzennego. Temu z kolei odpowiada ściśle określone rozsunięcie impulsów nadmiarowych elektronów i nadmiarowych dziur.

Przemieszczanie się impulsu nadmiarowych nośników większościowych (w rozważanym tu przypadku – elektronów) w kierunku przeciwnym do kierunku przepływu tych nośników jest spowodowane różnicą ich prędkości w obszarze impulsu i poza tym obszarem. Poza obszarem impulsu elektrony poruszają się z większą prędkością, a zatem dochodząc do „czoła" impulsu zostają przyhamowane, co sprawia, że impuls nadmiarowych elektronów „nadbudowuje się"

[1] Nie bierzemy tu pod uwagę obszarów ładunku przestrzennego, w których w miarę przesuwania się impulsu nośników nadmiarowych zmienia się natężenie pola elektrycznego, a zatem warunek niezależności prądu unoszenia od współrzędnej x nie jest spełniony.

od strony dopływu elektronów[1]. Odwrotnie zachowują się elektrony na drugim końcu impulsu. Tu powiększają one swą prędkość, co wywołuje zmniejszenie się koncentracji elektronów i w wyniku tego, przesuwanie się „tylnego zbocza" impulsu w kierunku przeciwnym do kierunku przepływu elektronów.

Tak więc impulsy nadmiarowych elektronów i nadmiarowych dziur przesuwają się razem, złączone potężnymi siłami elektrycznymi, które umożliwiają tylko takie ich rozsunięcie, jakie jest niezbędne do spełnienia warunku (12/40). Kierunek przesuwania się tego ambipolarnego impulsu jest uwarunkowany znakiem ładunku nośników *mniejszościowych*, a prędkość tego przesuwania w dostatecznie silnie domieszkowanym półprzewodniku – ruchliwością nośników mniejszościowych.

Zmniejszenie koncentracji domieszek, przy zachowaniu niezmienionych koncentracji nadmiarowych, sprawia, że ΔE rośnie, a wskutek tego ruchliwość grupowa maleje. Impuls nadmiarowych nośników przesuwa się więc z coraz mniejszą prędkością, a dla półprzewodnika samoistnego prędkość ta staje się równa zeru. Wynika to z analizy przeprowadzonej w problemie 12.2, jak również ze wzoru (12/28). Dla półprzewodnika samoistnego mamy bowiem $n = n_i + \Delta n = n_i + \Delta p = p$, a więc ruchliwość grupowa μ^* jest równa zeru.

Po zmianie typu przewodnictwa, tzn. po wprowadzeniu domieszek akceptorowych, impuls nośników nadmiarowych porusza się w kierunku przeciwnym niż poprzednio. Jest on więc teraz unoszony w kierunku zgodnym z kierunkiem ruchu elektronów, przy czym przy dostatecznie dużej koncentracji akceptorów ruchliwość grupowa osiąga wartość zgodną ze wzorem (12/34).

W powyższej analizie uwzględniliśmy tylko zjawisko unoszenia nośników nadmiarowych w polu elektrycznym. Pominęliśmy natomiast ich dyfuzję i rekombinację. Rozważania te są więc słuszne w warunkach, w których współczynniki dyfuzji nośników są dostatecznie małe (co może występować w bardzo niskich temperaturach), a ich czasy życia – dostatecznie długie. W przeciwnym razie impuls nośników nadmiarowych ulega stopniowemu rozmyciu w wyniku dyfuzji nośników, a liczba nośników nadmiarowych – stopniowemu zmniejszaniu się w wyniku ich rekombinacji. Przy małym poziomie zakłócenia wszystkie wymienione tu procesy zachodzą niezależnie od siebie, a zatem każdy z nich może być rozpatrywany osobno.

W celu zilustrowania tego stwierdzenia rozważmy analizowany poprzednio przypadek, przyjmując, że impuls nośników nadmiarowych jest w chwili pojawienia się nieskończenie wąski. Dla półprzewodnika typu n rozwiązanie ambipolarnego równania transportu ma dla $t > 0$ następującą postać:

$$\Delta p\,(x, t) = \frac{a}{\sqrt{D_p t}} \exp\left[-\frac{(x-\mu_p E_0 t)^2}{4D_p t}\right] \exp\left(-\frac{t}{\tau_p}\right) \qquad (12/41)$$

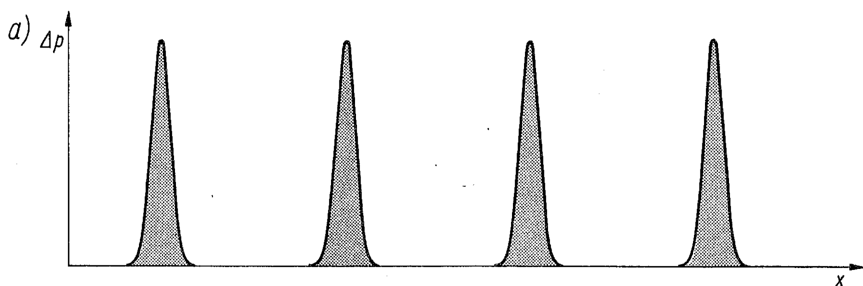
gdzie a jest stałą zależną od początkowej liczby nośników nadmiarowych.

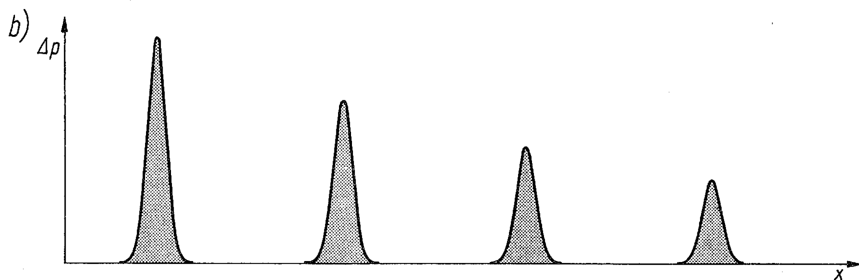
[1] Podobne zjawisko można obserwować np. na autostradzie, gdy na pewnym jej odcinku powstaje zator. Samochody dojeżdżające do zablokowanego odcinka autostrady muszą zwolnić, wskutek czego koniec zatoru przesuwa się w kierunku przeciwnym do kierunku ruchu samochodów.

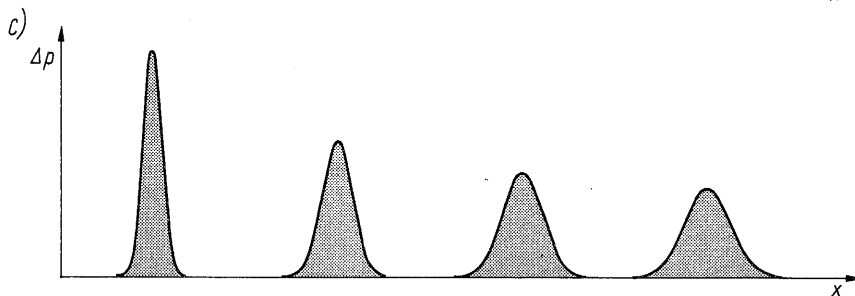
Impuls nośników nadmiarowych ma więc kształt gaussowski. Jego maksimum zmienia swe położenie zgodnie z równaniem

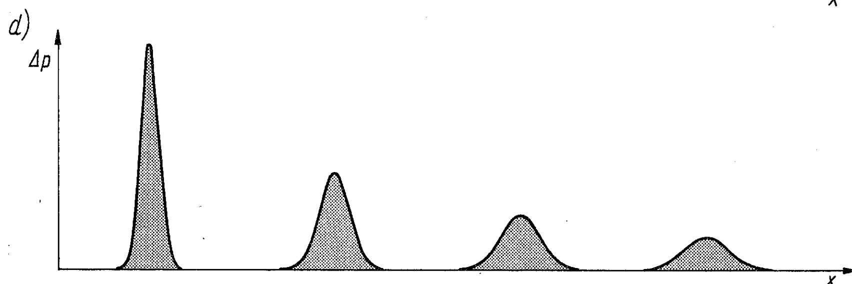
$$x-\mu_p Et = 0 \tag{12/42}$$

Przesuwa się ono więc w kierunku zgodnym z kierunkiem pola elektrycznego z prędkością równą $\mu_p E_0$ – jest to efekt *unoszenia* nośników przez pole.









12/10 Zachowanie się impulsu nośników nadmiarowych w polu elektrycznym w różnych warunkach: a) brak dyfuzji i rekombinacji; b) brak dyfuzji; c) brak rekombinacji; d) warunki rzeczywiste

W miarę przesuwania się impulsu jego szerokość zwiększa się. Dyspersja rozkładu wynosi bowiem

$$2D_pt$$

a więc wzrasta proporcjonalnie do czasu mierzonego od chwili wytworzenia impulsu – efekt *dyfuzji* nośników.

Całkowita liczba nośników nadmiarowych jest proporcjonalna do wyrażenia

$$\exp\left(-\frac{t}{\tau_p}\right)$$

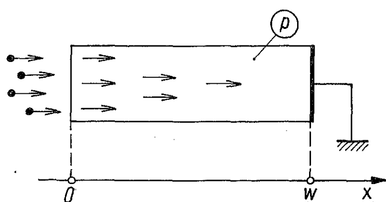
w miarę upływu czasu zanika ona więc wykładniczo ze stałą czasową równą τ_p – efekt *rekombinacji* nośników.

Zachowanie się gaussowskiego impulsu nośników nadmiarowych w różnych warunkach pokazano na rys. 12/10.

12.5 PRZYKŁADY ROZWIĄZANIA AMBIPOLARNEGO RÓWNANIA TRANSPORTU

12.5.1 Wstrzykiwanie nośników mniejszościowych

Rozpatrzmy przedstawioną na rys. 12/11 próbkę półprzewodnika typu p zakończoną idealnym kontaktem rekombinacyjnym na potencjale ziemi. Do próbki tej jest wprowadzany z zewnątrz strumień nośników mniejszościowych (elektronów). Nośniki te dyfundują w głąb półprzewodnika i stopniowo zanikają wskutek rekombinacji w obszarze półprzewodnika lub na powierzchni kontaktu[1]. W stanie ustalonym występuje w półprzewodniku określony rozkład nadmiarowych elektronów $\Delta n(x)$ i, zgodnie z warunkiem quasi-neutralności, prawie identyczny rozkład nadmiarowych dziur $\Delta p(x)$.



12/11 Wstrzykiwanie nośników mniejszościowych do próbki półprzewodnika typu *p*

Zadaniem naszym będzie wyznaczenie tych rozkładów. Ograniczymy się przy tym do małego poziomu zakłócenia (który w tym przypadku nazywamy *małym poziomem wstrzykiwania*), gdyż tylko przy tym założeniu problem daje się rozwiązać analitycznie. Zgodnie z tym założeniem, możemy skorzystać z ambipolarnego równania transportu o postaci (12/37).

---

[1] Jak zaznaczono na początku tego rozdziału, rozpatrujemy tu jedynie przepływy jednowymiarowe, nie bierzemy więc pod uwagę możliwości dopływu elektronów do powierzchni bocznych próbki. Założenie jednowymiarowości jest usprawiedliwione m.in. tym, że w większości elementów półprzewodnikowych rozmiary poprzedzane obszaru, do którego wstrzykiwane są nośniki, są wielokrotnie większe niż długość tego obszaru.

Ponieważ rozpatrujemy stan ustalony, zatem prawą stronę równania można przyrównać do zera. Można również położyć $g = 0$ (półprzewodnik nie jest oświetlony promieniowaniem wywołującym generację par elektron-dziura) oraz $E = 0$ (zgodnie z założeniem małego poziomu wstrzykiwania pomijamy bowiem wpływ pola elektrycznego na zachowanie się nośników *mniejszościowych*). Ostatecznie więc ambipolarne równanie transportu sprowadza się do postaci

$$D_n \frac{d^2\Delta n}{dx^2} - \frac{\Delta n}{\tau_n} = 0 \tag{12/43}$$

a jego rozwiązaniem ogólnym jest funkcja

$$\Delta n(x) = A \exp\left(-\frac{x}{L_n}\right) + B \exp\left(\frac{x}{L_n}\right) \tag{12/44}$$

przy czym wprowadzony tu nowy parametr L_n jest zdefiniowany wzorem

$$\boxed{L_n = \sqrt{D_n \tau_n}} \tag{12/45}$$

Funkcja (12/44) musi spełniać warunki brzegowe stosowne dla rozpatrywanego zagadnienia. Dotyczą one koncentracji nadmiarowych elektronów na początku i na końcu próbki. Warunki te brzmią następująco:

$$\Delta n(x)|_{x=0} = \Delta n(0) \tag{12/46}$$

$$\Delta n(x)|_{x=w} = 0 \tag{12/47}$$

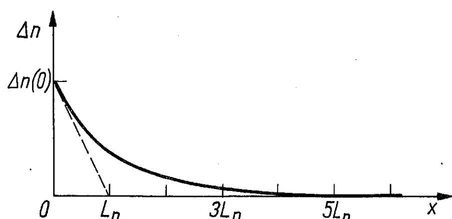
Ten ostatni warunek wynika z założenia, że próbka jest zakończona idealnym uziemionym kontaktem rekombinacyjnym.

Przyjmijmy początkowo, że rozpatrywana próbka jest na tyle długa, że koncentracja nadmiarowych elektronów maleje wskutek rekombinacji praktycznie do zera zanim elektrony te osiągną kontakt znajdujący się na końcu próbki. Obszar półprzewodnika spełniający ten warunek nazywamy *bazą długą*. Dla takiej bazy stałe dowolne w wyrażeniu (12/44) przyjmują następujące wartości: $A = \Delta n(0)$; $B = 0$. Funkcja opisująca rozkład koncentracji nadmiarowych elektronów przybiera zatem postać

$$\Delta n(x) = \Delta n(0) \exp\left(-\frac{x}{L_n}\right) \tag{12/48}$$

Wykres tej funkcji przedstawiono na rys. 12/12. Jak widzimy, koncentracja nadmiarowych elektronów maleje wzdłuż próbki wykładniczo, przy czym wielkość L_n jest stałą charakteryzującą szybkość zaniku nośników nadmiarowych w funkcji odległości, podobnie jak czas życia nośników nadmiarowych jest stałą charakteryzującą szybkość zaniku tych nośników w funkcji czasu.

Wielkość L_n zdefiniowaną wzorem (12/45) nazywamy *drogą dyfuzji* elektronów w półprzewodniku typu *p*. Analogicznie definiuje się drogę dyfuzji dziur w półprzewodniku typu *n*.



12/12 Rozkład koncentracji nadmiarowych nośników mniejszościowych w przypadku przedstawionym na rys. 12/11

Jak łatwo można sprawdzić, droga dyfuzji nośników mniejszościowych jest to średnia odległość przebywana przez nośniki mniejszościowe wstrzykiwane do próbki półprzewodnika o dostatecznie dużej długości.

Ze wzoru (12/48) wynika, że na odcinku o długości L_n koncentracja nośników nadmiarowych maleje e-krotnie, a na odcinku o długości $5L_n$ spada do wartości mniejszej niż 1% wartości na początku próbki. Przyjęte założenie, że rozpatrywana próbka półprzewodnika jest tak długa, że wprowadzone do niej elektrony praktycznie nie docierają do kontaktu rekombinacyjnego, można więc wyrazić za pomocą nierówności

$$w > 5L_n \tag{12/49}$$

Dla krzemu o niezbyt dużej koncentracji domieszek mamy w temperaturze pokojowej: $D_n = 36\ \text{cm}^2/\text{s}$, a $D_p = 12\ \text{cm}^2/\text{s}$. Przyjmując, że czas życia nośników nadmiarowych jest w obu przypadkach równy 1 μs, otrzymujemy następujące orientacyjne wartości drogi dyfuzji nośników mniejszościowych:

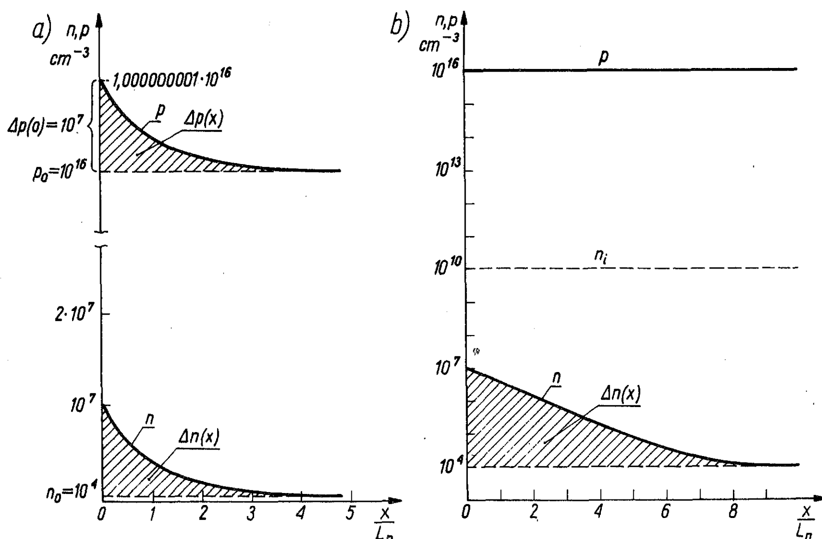
$$L_n = 60\ \mu\text{m}; \qquad L_p = 35\ \mu\text{m}$$

Przy większych koncentracjach domieszek maleją zarówno ruchliwości nośników, jak i ich czasy życia, a zatem droga dyfuzji nośników mniejszościowych zmniejsza się do wartości kilkunastu, a nawet kilku mikrometrów.

Znajomość rozkładu koncentracji nadmiarowych nośników mniejszościowych – wzór (12/48) – umożliwia wyznaczenie rozkładów całkowitych koncentracji elektronów i dziur w rozpatrywanej próbce. Przykład takich rozkładów podano na rys. 12/13. Dotyczy on półprzewodnika typu p o koncentracji akceptorów $N_a = 10^{16}\ \text{cm}^{-3}$. Koncentracje równowagowe dziur i elektronów wynoszą więc odpowiednio $p_0 = 10^{16}\ \text{cm}^{-3}$ i $n_0 = 10^4\ \text{cm}^{-3}$. W przedstawionym przykładzie przyjęto, że $\Delta n(0) = 10^7\ \text{cm}^{-3}$.

W lewej części rysunku pokazano wykresy $n(x)$ i $p(x)$ przy zastosowaniu na osi rzędnych skali liniowej, natomiast w prawej części rysunku – te same wykresy w skali logarytmicznej. W skali liniowej są widoczne bezwzględne odchylenia koncentracji nośników od wartości równowagowych. Ze względu na warunek quasi-neutralności, odchylenia te są dla elektronów i dziur jednakowe – wzór (12/25).

Przy zastosowaniu skali logarytmicznej uwidaczniają się względne odchylenia koncentracji nośników od ich wartości równowagowych. W przypadku



12/13 Rozkład elektronów i dziur w przypadku przedstawionym na rys. 12/11; skala na osi rzędnych liniowa (a) i logarytmiczna (b)

elektronów są one bardzo znaczne, natomiast w przypadku dziur – niedostrzegalnie małe. Jest to spowodowane ograniczeniem naszych rozważań do małego poziomu wstrzykiwania. Możemy więc stwierdzić, że mały poziom zakłócenia jest to taki stan nierównowagowy, w którym koncentracja nośników *mniejszościowych* ulega istotnemu zakłóceniu, natomiast koncentracja nośników *większościowych* pozostaje praktycznie niezakłócona.

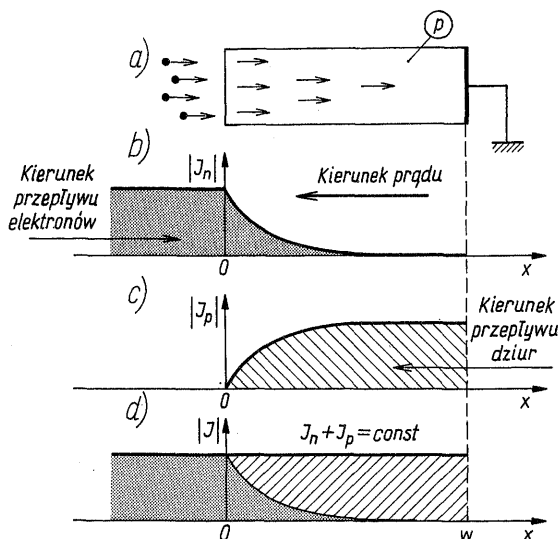
Dyfuzja elektronów wzdłuż rozważanej próbki (rys. 12/11) wywołuje przepływ prądu elektrycznego w próbce. Gęstość tego prądu można wyznaczyć na podstawie wzoru (10/61). Ponieważ założyliśmy, że przy małym poziomie wstrzykiwania wpływ pola elektrycznego na zachowanie się nośników *mniejszościowych* jest pomijalnie mały, zatem przy wyznaczaniu gęstości prądu elektronowego możemy ograniczyć się do składnika dyfuzyjnego. Otrzymujemy wówczas

$$J_n(x) = qD_n \frac{\mathrm{d}n}{\mathrm{d}x} = -\frac{qD_n}{L_n} \Delta n(0) \exp\left(-\frac{x}{L_n}\right) \tag{12/50}$$

Prąd płynie w kierunku przeciwnym do kierunku osi x, a jego wartość zanika wykładniczo wzdłuż próbki – zob. rys. 12/14b. Największa wartość prądu występuje w punkcie $x = 0$. Jest ona równa prądowi wywołanemu strumieniem elektronów dolatujących do próbki.

Z porównania wzorów (12/50) i (12/48) wynika, że gęstość prądu elektronowego w danym punkcie jest proporcjonalna do koncentracji nadmiarowych elektronów w tym punkcie. Oznacza to, że przy wykładniczym rozkładzie nośników nadmiarowych w próbce, dyfuzja nośników mniejszościowych zachodzi ze stałą prędkością

$$v_{nd} = \frac{D_n}{L_n} \tag{12/51}$$



12/14 Prądy w półprzewodniku w przypadku przedstawionym na rys. 12/11: a) próbka półprzewodnika; b) prąd elektronowy; c) prąd dziurowy; d) prąd całkowity

Wniosek ten wynika również bezpośrednio ze wzoru (10/60).

Wykładniczy zanik koncentracji nadmiarowych elektronów $\Delta n\,(x)$ i gęstości prądu elektronowego $J_n(x)$ jest spowodowany rekombinacją tych nośników z dziurami, których koncentracja w półprzewodniku typu p jest bardzo duża. Dziury znikające w wyniku rekombinacji zostają uzupełnione dziurami wytwarzanymi na powierzchni kontaktu rekombinacyjnego. Wzdłuż próbki płynie zatem prąd dziurowy, którego wartość jest największa w punkcie $x = w$ i maleje monotonicznie do zera w punkcie $x = 0$ – zob. rys. 12/14c.

Prądu tego nie można wyznaczyć ze wzoru (10/65), ponieważ nie jest znany rozkład pola elektrycznego wzdłuż próbki. Wpływ tego pola można było zaniedbać przy analizowaniu zachowania się nośników mniejszościowych, ponieważ przy małym poziomie wstrzykiwania koncentracje tych nośników są niewielkie. W przypadku nośników większościowych, których koncentracje są duże, wpływ pola elektrycznego nie może być zaniedbany. Jak zobaczymy w następnym punkcie, prąd wywołany unoszeniem dziur jest porównywalny z prądem dyfuzyjnym.

Nie mogąc skorzystać ze wzoru (10/65), musimy wyznaczyć prąd dziurowy w inny sposób. Ponieważ w stanie ustalonym prąd w dowolnym przekroju obwodu elektrycznego ma jednakową wartość, zatem całkowita gęstość prądu jest stała i wynosi

$$J = J_n(x)+J_p(x) = -q\frac{D_n}{L_n}\Delta n\,(0) \qquad (12/52)$$

Ze wzoru tego otrzymuje się

$$J_p(x) = -q\frac{D_n}{L_n}\Delta n\,(0)\left[1-\exp\left(-\frac{x}{L_n}\right)\right] \qquad (12/53)$$

Problem 12.3. Rozważ zjawisko wstrzykiwania elektronów do próbki półprzewodnika typu *p* (rys. 12/11), zakładając, że próbka ta spełnia warunek (12/49). Wyznacz związek zachodzący między ładunkiem nadmiarowych nośników mniejszościowych a prądem płynącym przez próbkę. Podaj przykład innego zjawiska, w którym zachodzi związek o podobnym charakterze.

Jeżeli warunek $w \gg L_n$ nie jest spełniony, to rozwiązanie równania (12/43) jest nieco bardziej skomplikowane. Rozwiązanie ogólne może być przedstawione w postaci funkcji (12/44), można jednakże użyć dowolnej innej kombinacji funkcji wykładniczych o wykładnikach x/L_n i $-x/L_n$. W rozpatrywanym przypadku dogodnie jest przyjąć, że rozwiązanie ogólne ma postać

$$\Delta n(x) = A \operatorname{sh} \frac{x}{L_n} + B \operatorname{ch} \frac{x}{L_n} \tag{12/54}$$

Zastosowanie warunków brzegowych (12/46) i (12/47) prowadzi do wniosku, że

$$A = -\Delta n(0) \operatorname{cth} \frac{w}{L_n}; \qquad B = \Delta n(0) \tag{12/55}$$

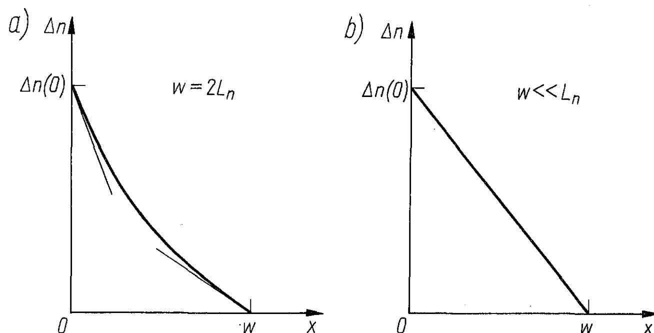
co po podstawieniu do (12/45) daje

$$\Delta n(x) = \frac{\operatorname{sh} \dfrac{w-x}{L_n}}{\operatorname{sh} \dfrac{w}{L_n}} \Delta n(0) \tag{12/56}$$

Wykres tej funkcji przedstawiono na rys. 12/15. Dla bazy długiej (tzn. dla $w \gg L_n$) wyrażenie (12/56) upraszcza się, jak łatwo można sprawdzić, do postaci (12/48), natomiast dla tzw. *bazy krótkiej* (tzn. dla $w \ll L_n$), otrzymujemy

$$\Delta n(x) \approx \Delta n(0)\left(1 - \frac{x}{L_n}\right) \tag{12/57}$$

Koncentracja nośników nadmiarowych maleje więc w przybliżeniu liniowo wzdłuż próbki – zob. rys. 12/15b. Fakt ten świadczy o tym, że gęstość prądu



12/15 Rozkład koncentracji nośników nadmiarowych w próbkach o różnej wartości stosunku w/L_n: a) $w = 2L_n$; b) $w \ll L_n$

elektronowego jest prawie jednakowa w każdym przekroju próbki[1]. Inaczej mówiąc, niemal wszystkie elektrony wstrzykiwane do próbki dochodzą do kontaktu rekombinacyjnego, a w obszarze bazy prawie nie zachodzi rekombinacja nośników.

Baza długa i baza krótka stanowią 2 krańcowe przypadki. W pierwszym, praktycznie rzecz biorąc, elektrony nie docierają do kontaktu rekombinacyjnego, ulegają bowiem rekombinacji w obszarze bazy. O obszarze takim mówimy, że jest on dla elektronów całkowicie *„nieprzezroczysty"*. W drugim przypadku natomiast docierają one do kontaktu rekombinacyjnego prawie bez strat spowodowanych rekombinacją elektronów we wnętrzu półprzewodnika – jest to więc obszar całkowicie *„przezroczysty"* dla elektronów.

W ogólnym przypadku część elektronów ulega rekombinacji w obszarze bazy, pozostała część zaś dociera do kontaktu rekombinacyjnego. Można to odczytać z rys. 12/15a, stwierdzając, że gęstość prądu elektronowego na końcu bazy jest mniejsza niż na jej początku.

Problem 12.4. Dla bazy o długości $w = 2L_n$ wykonaj wykresy analogiczne do przedstawionych na rys. 12/14. Przedyskutuj otrzymane wyniki.

Problem 12.5. W celu scharakteryzowania liczby elektronów docierających do kontaktu rekombinacyjnego, wprowadza się pojęcie *współczynnika sprawności transportu* elektronów przez bazę, określając go wzorem

$$\alpha_T = \frac{I_n(w)}{I_n(0)}$$

gdzie $I_n(x)$ oznacza wartość prądu elektronowego w przekroju x. Wyprowadź ogólny wzór na współczynnik α_T. Przedstaw za pomocą wykresów zależności $I_n(0)$, $I_n(w)$ i α_T od stosunku w/L_n. Przedyskutuj otrzymane wyniki.

Problem 12.6. Wyprowadź wzór na czas przelotu nadmiarowych elektronów wstrzykiwanych do półprzewodnika typu p przy założeniu, że $w \ll L_n$. Opierając się na tym wzorze, znajdź związek zachodzący między ładunkiem nadmiarowych nośników mniejszościowych w bazie krótkiej a płynącym przez nią prądem. Porównaj otrzymany wzór z wynikiem uzyskanym w probl. 12.3.

Wskazówka: Czas przelotu elektronu przez rozpatrywany obszar jest to czas zużywany na przebycie drogi od punktu $x = 0$ do $x = w$. Czas ten można wyznaczyć ze wzoru

$$\tau = \int_0^w \frac{dx}{v(x)}$$

gdzie $v(x)$ oznacza prędkość elektronu w punkcie x. Tę ostatnią można z kolei wyrazić za pomocą wzoru (10/60).

Problem 12.7. Zjawiskiem odwrotnym do wstrzykiwania nośników mniejszościowych do danego obszaru jest ich wyciąganie (ekstrakcja) z tego obszaru. Rozważ to zjawisko dla próbki przedstawionej na rys. 12/11, zakładając, że $w \gg L_n$. Przyjmując, że koncentracja elektronów w płaszczyźnie $x = 0$ jest równa zeru, wyznacz:

– rozkład koncentracji elektronów $n(x)$;
– rozkłady gęstości prądów $J_n(x)$ i $J_p(x)$.

---

[1] Wobec pominięcia działania pola elektrycznego, gęstość prądu elektronowego jest proporcjonalna do gradientu koncentracji elektronów. Przy liniowej zależności $\Delta n(x)$ gradient ten ma wartość stałą, a zatem gęstość prądu elektronowego nie zależy od współrzędnej x.

Wykaż, że liczba elektronów wypływających w jednostce czasu z rozpatrywanej próbki półprzewodnika jest równa liczbie elektronów pojawiających się w wyniku generacji termicznej w obszarze o grubości równej L_n.

12.5.2* Analiza warunku quasi-neutralności

Rozważania zawarte w poprzednim punkcie były oparte na założeniu, że warunek quasi-neutralności jest spełniony z bardzo dobrym przybliżeniem, w związku z czym wpływ pola elektrycznego na zachowanie się nośników mniejszościowych jest pomijalnie mały. Obecnie przystąpimy do zbadania, w jakim stopniu założenie to jest usprawiedliwione. Rozważania nasze będą dotyczyły przypadku bazy długiej, a metoda postępowania będzie następująca: w celu wyznaczenia natężenia pola elektrycznego należy wyznaczyć prąd unoszenia dziur, a następnie skorzystać ze związku zachodzącego między J_{pu} i E.

Gęstość prądu dziurowego w rozważanym przypadku wyraża się wzorem (12/53), który można również przedstawić w następującej postaci:

$$J_p(x) = J - J_n(x) \tag{12/58}$$

gdzie: J – całkowita gęstość prądu, a $J_n(x)$ – gęstość prądu elektronowego w danym punkcie. Ta ostatnia, zgodnie z przyjętym założeniem, zawiera jedynie składnik dyfuzyjny – zob. wzór (12/50). Prąd dziurowy musi być natomiast traktowany jako suma odpowiedniego prądu dyfuzji i prądu unoszenia. Gęstość prądu dyfuzyjnego wynosi przy tym

$$J_{pd}(x) = -qD_p \frac{\mathrm{d}p}{\mathrm{d}x} = -\frac{1}{b} J_n(x) \tag{12/59}$$

gdzie b jest to współczynnik równy stosunkowi ruchliwości μ_n do μ_p dla danego półprzewodnika (przykładowo dla krzemu b wynosi ok. 3). Ostatnia równość we wzorze (12/59) wynika z założenia, że $\Delta p = \Delta n$.

Ze wzorów (12/58) i (12/59) wynika, że gęstość prądu unoszenia dziur wynosi

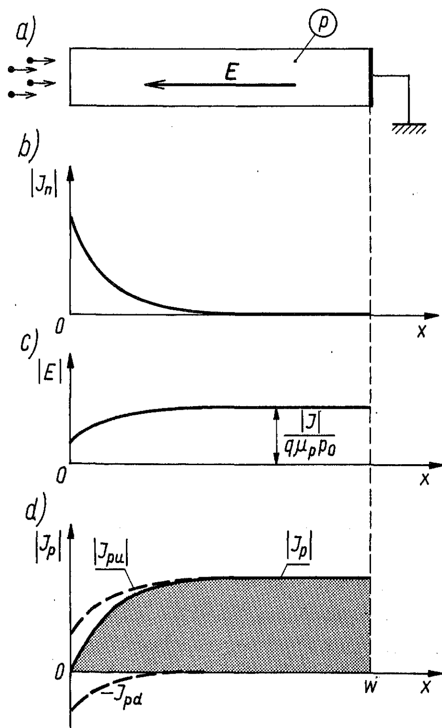
$$J_{pu}(x) = J_p(x) - J_{pd}(x) = J - J_n(x)\left(1 - \frac{1}{b}\right) \tag{12/60}$$

Znając wartość tego prądu, otrzymujemy następujące wyrażenie na natężenie pola elektrycznego:

$$E(x) = \frac{J_{pu}(x)}{q\mu_p p_0} = \frac{J - J_n(x)\left(1 - \frac{1}{b}\right)}{q\mu_p p_0} \tag{12/61}$$

Jako koncentrację dziur podstawiliśmy tu wartość równowagową p_0, ponieważ przy małym poziomie wstrzykiwania koncentracja nośników większościowych praktycznie nie ulega zakłóceniu.

Pole elektryczne jest skierowane zgodnie z kierunkiem przepływu prądu w próbce, a więc przeciwnie do kierunku osi x – zob. rys. 12/16a. Wstrzykiwanie nośników do półprzewodnika wymaga więc wytworzenia różnicy poten-



12/16
Wielkości związane z przepływem nośników w przypadku przedstawionym na rys. 12/11: a) próbka półprzewodnika; b) prąd elektronowy; c) natężenie pola elektrycznego; d) prąd dziurowy

cjałów między płaszczyzną, przez którą nośniki są wprowadzane do półprzewodnika, a kontaktem rekombinacyjnym. Wytworzone w wyniku tego pole elektryczne utrzymuje przedstawiony na rys. 12/13a rozkład koncentracji dziur odbiegających od rozkładu równowagowego i zapewnia przepływ tych nośników wzdłuż próbki półprzewodnika.

Rozkład tego pola w rozpatrywanej próbce przedstawiono na rys. 12/16c. Przy równości ruchliwości μ_n i μ_p pole to byłoby jednorodne, natomiast przy $b \neq 1$ jest ono jednorodne dopiero w obszarze, w którym praktycznie nie występują nośniki nadmiarowe.

W dolnej części rys. 12/16 pokazano zależności gęstości prądu dziurowego oraz jego składników J_{pu} i J_{pd} od współrzędnej x. Jak widzimy, na początku próbki wartości tych składników są porównywalne, natomiast w dalszej części próbki składnik dyfuzyjny przestaje odgrywać istotną rolę. Prąd dziurowy jest tu wywołany prawie wyłącznie unoszeniem dziur w polu elektrycznym.

Znajomość rozkładu pola elektrycznego pozwala ocenić, jaką część prądu płynącego przez próbkę stanowi pominięty przez nas prąd unoszenia elektronów. Największa wartość tego prądu występuje tam, gdzie koncentracja tych nośników jest największa, tzn. na początku próbki. Z wyprowadzonych wcześniej wzorów wynika, że

$$\frac{J_{nu}(0)}{J} = \frac{q\mu_n n(0) E(0)}{J} = \frac{n(0)}{p_0} \tag{12/62}$$

Przy małym poziomie wstrzykiwania $n(0) \ll p_0$, a więc istotnie prąd unoszenia nośników mniejszościowych stanowi pomijalnie małą część całkowitego prądu przepływającego przez rozpatrywaną próbkę.

Problem 12.8. Wykonaj wykresy analogiczne do przedstawionych na rys. 12/16 dla próbki półprzewodnika typu n, do której są wstrzykiwane dziury. Przyjmij $b = 3$.

Warto jeszcze sprawdzić, z jaką dokładnością jest spełniony warunek quasi-neutralności w rozważanym tu przypadku. W tym celu posłużymy się równaniem (11/10b), które może być przepisane w postaci

$$\Delta p - \Delta n = \frac{\varepsilon}{q} \frac{\mathrm{d}E}{\mathrm{d}x} \tag{12/63}$$

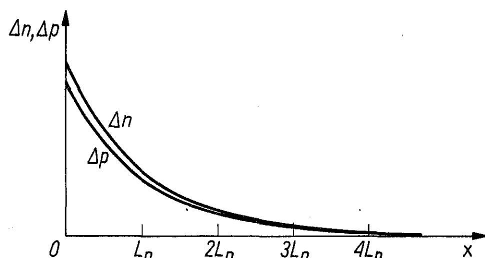
Korzystając z zależności (12/61), otrzymamy po pewnych przekształceniach

$$\frac{\Delta n - \Delta p}{\Delta n} = \left(1 - \frac{1}{b}\right)\frac{\tau_d}{\tau_n} \tag{12/64}$$

gdzie: τ_d – czas relaksacji dielektrycznej, a τ_n – czas życia nadmiarowych elektronów. Przyjmując przykładowo, że τ_d wynosi 1 ps, a τ_n – 1 μs, widzimy, że koncentracje nadmiarowych elektronów i nadmiarowych dziur różnią się zaledwie o jedną milionową. Stwierdziliśmy więc, tak jak to już wcześniej zaznaczono, że warunek quasi-neutralności elektrycznej jest spełniony z bardzo dużą dokładnością.

Wyrażenie występujące po prawej stronie równości (12/64) jest niezależne od współrzędnej x, a zatem względne odchylenie Δp od Δn jest jednakowe w całej rozpatrywanej próbce. Rozkład koncentracji nośników nadmiarowych wygląda więc orientacyjnie tak, jak to pokazano na rys. 12/17. W celu umożliwienia wykonania tego rysunku, powiększono na nim różnicę $\Delta n - \Delta p$ 100 000 razy.

Problem 12.9. Wyprowadź wzór (12/64).



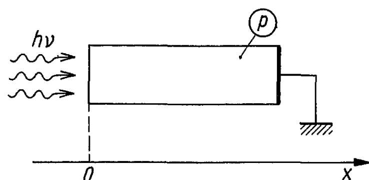
12/17 Rozkład koncentracji nadmiarowych elektronów i dziur w przypadku przedstawionym na rys. 12/11

12.5.3* Dyfuzja dwuskładnikowa

W ostatnich 2 punktach rozpatrywaliśmy dyfuzję nadmiarowych elektronów wprowadzanych z zewnętrznego źródła do półprzewodnika typu p. Stwierdziliśmy tam że przemieszczaniu się elektronów w kierunku kontaktu rekombinacyjnego towarzyszy przepływ dziur w przeciwnym kierunku. Przepływ ten jest

wywołany polem elektrycznym, które musi być w tym celu wytworzone w półprzewodniku.

Obecnie zajmiemy się przypadkiem, w którym nośniki obu rodzajów (elektrony i dziury) poruszają się w równych ilościach w tym samym kierunku. Zjawisko to nazywamy *dyfuzją dwuskładnikową* lub *dyfuzją ambipolarną*. Rozpatrzymy je na przykładzie prostopadłościennej próbki półprzewodnika zakończonej idealnym kontaktem rekombinacyjnym – rys. 12/18.



12/18 Próbka półprzewodnika poddana działaniu promieniowania elektromagnetycznego

Na powierzchnię próbki pada promieniowanie elektromagnetyczne wywołujące generację par elektron-dziura. W celu uproszczenia rozważań, załóżmy, że promieniowanie to jest pochłaniane w nieskończenie cienkiej warstwie przypowierzchniowej, a więc we wnętrzu próbki szybkość generacji nośników nadmiarowych jest równa zeru.

Nośniki nadmiarowe wytwarzane w warstwie przypowierzchniowej dyfundują w głąb półprzewodnika i stopniowo zanikają wskutek rekombinacji we wnętrzu półprzewodnika i na powierzchni kontaktu. Jeżeli rozpatrywany półprzewodnik jest typu p, to w stanie ustalonym rozkład koncentracji nośników nadmiarowych jest opisany tym samym równaniem (12/43), którym posługiwaliśmy się przy rozpatrywaniu wstrzykiwania nośników mniejszościowych. Ponieważ warunki brzegowe są także identyczne, zatem rozwiązanie tego równania będzie miało postać zgodną ze wzorem (12/56), wyprowadzonym w p. 12.5.1.

Na tym kończą się jednakże podobieństwa tych dwóch zjawisk. Zachowanie się nośników jest bowiem w obu przypadkach zupełnie różne. Przy dyfuzji dwuskładnikowej nośniki nadmiarowe obu rodzajów poruszają się w kierunku malejącej koncentracji, tzn. od powierzchni oświetlonej do kontaktu rekombinacyjnego. Współczynniki dyfuzji tych nośników są jednakże różne, przy czym z reguły ruchliwość elektronów μ_n jest większa niż ruchliwość dziur μ_p, a co za tym idzie również $D_n > D_p$. Ta różnica ruchliwości sprawia, że początkowo ilość elektronów dyfundujących w kierunku kontaktu rekombinacyjnego przewyższa analogiczną ilość dziur. W objętości półprzewodnika pojawia się więc nadmiar ładunku ujemnego, a w bardzo cienkiej warstwie w pobliżu oświetlonej powierzchni – nadmiar ładunku przeciwnego znaku. W próbce półprzewodnika powstaje więc pole elektryczne skierowane zgodnie z kierunkiem osi x. Hamuje ono ruch nośników o większej ruchliwości (tzn. elektronów), a przyspiesza nośniki mniej ruchliwe (dziury).

Stan ustalony powstaje przy takim rozkładzie ładunku przestrzennego i związanego z nim pola elektrycznego, przy którym w każdym przekroju próbki

ilości przemieszczających się elektronów i dziur są jednakowe. W tych warunkach gęstość prądu elektronowego jest co do wartości bezwzględnej równa gęstości prądu dziurowego, a prąd występujący w półprzewodniku staje się równy zeru.

Z warunku $J_n + J_p = 0$ można wyznaczyć rozkład natężenia pola elektrycznego w półprzewodniku. Korzystając ze wzorów (10/64) i (10/65) oraz z warunku quasi-neutralności, otrzymuje się

$$E(x) = \frac{D_p - D_n}{n\mu_n + p\mu_p} \frac{\mathrm{d}n}{\mathrm{d}x} \tag{12/65}$$

Ze wzoru tego wynika, że natężenie pola ma wartość dodatnią, a zatem, jak to już powiedziano, pole elektryczne w półprzewodniku jest skierowane zgodnie z kierunkiem osi x. Wytwarzanie się pola elektrycznego w jednorodnym półprzewodniku w wyniku absorpcji promieniowania elektromagnetycznego nosi nazwę *zjawiska Dembera*. Jak wynika z przeprowadzonej analizy, przyczyną tego zjawiska jest nierówność współczynnika dyfuzji D_n i D_p, a uzyskany efekt jest proporcjonalny do różnicy $D_n - D_p$.

Podstawiając wyrażenie (12/65) do wzoru (10/64), znajdziemy, że gęstość prądu elektronowego w rozpatrywanej próbce wynosi

$$J_n = q \frac{n+p}{\dfrac{n}{D_p} + \dfrac{p}{D_n}} \frac{\mathrm{d}n}{\mathrm{d}x} \tag{12/66}$$

Występujące w tym wzorze wyrażenie ułamkowe jest to ambipolarny współczynnik dyfuzji D^* określony wzorem (12/29). Tak więc przepływ nośników obu rodzajów można opisać za pomocą współczynnika D^*, w którym ukryte jest działanie pola elektrycznego. Współczynnik ten przyjmuje wartości pośrednią między D_n a D_p.

Problem 12.10. Na powierzchnię próbki półprzewodnika typu p (rys. 12/18) o koncentracji akceptorów N_a pada strumień fotonów wywołujących generację par elektron-dziura w bardzo cienkiej warstwie przypowierzchniowej. Przyjmując, że $w \gg L_n$ oraz że natężenie tego strumienia wyraża się liczbą S fotonów na cm^2 i sekundę, wyprowadź wzory na:

a) koncentrację nadmiarowych elektronów Δn w przekroju $x = 0$;

b) *napięcia Dembera*, tzn. różnicę potencjałów między oświetloną i nieoświetloną powierzchnią próbki.

Wyznacz te wielkości dla próbki krzemowej o następujących danych: $N_a = 10^{15}$ cm^{-3}, $S = 10^{17}$ cm$^{-2}\cdot$s^{-1}, $b = 3$, $\tau_n = 10\mu$s, $T = 300$ K.

DODATEK **A**

RÓWNANIA MAXWELLA

Równania Maxwella są to podstawowe równania elektrodynamiki klasycznej. Stanowią one syntezę i uogólnienie wszystkich uprzednio znanych praw dotyczących zjawisk elektrycznych i magnetycznych. Opisują one zachowanie się w czasie i przestrzeni wektorów natężenia pola elektrycznego $\boldsymbol{E}$, indukcji elektrycznej $\boldsymbol{D} = \varepsilon \boldsymbol{E}$, natężenia pola magnetycznego $\boldsymbol{H}$ i indukcji magnetycznej $\boldsymbol{B} = \mu \boldsymbol{H}$, gdy są znane rozkłady gęstości ładunku elektrycznego ϱ i gęstości prądu $\boldsymbol{J}$. Równania te mają następującą postać:

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{A/1}$$

$$\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J} \tag{A/2}$$

$$\nabla \cdot \boldsymbol{D} = \varrho \tag{A/3}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{A/4}$$

Z równań tych wynika, że w jednorodnym i izotropowym ośrodku przy braku prądów i ładunków, wektory pola elektromagnetycznego spełniają równania falowe[1]:

$$\nabla^2 \boldsymbol{E} - \varepsilon\mu \frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0 \tag{A/5}$$

$$\nabla^2 \boldsymbol{H} - \varepsilon\mu \frac{\partial^2 \boldsymbol{H}}{\partial t^2} = 0 \tag{A/6}$$

---

[1] Równania te otrzymuje się, wyznaczając rotacje wektorów $\nabla \times \boldsymbol{E}$ i $\nabla \times \boldsymbol{H}$ i korzystając z następującej zależności słusznej dla dowolnego pola wektorowego $\boldsymbol{A}$:

$$\nabla \times (\nabla \times \boldsymbol{A}) = \nabla (\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}$$

Równania (A/5) i (A/6) opisują rozchodzenie się fal elektromagnetycznych w danym ośrodku. Wynika z nich, że fale te rozchodzą się z prędkością określaną wzorem (1/26).

Spośród wielu wniosków, jakie można wyciągnąć z równań Maxwella, na szczególną uwagę zasługuje *równanie ciągłości*, wyrażające zasadę zachowania ładunków elektrycznych.

Ponieważ dla każdego pola wektorowego dywergencja rotacji jest równa zeru, zatem na mocy równania (A/2)

$$\frac{\partial}{\partial t}(\nabla \cdot \boldsymbol{D}) + \nabla \cdot \boldsymbol{J} = 0 \tag{A/7}$$

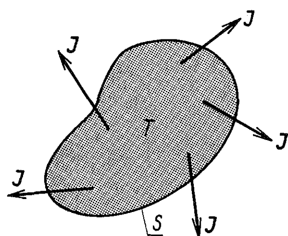
Korzystając teraz z równania (A/3), otrzymujemy równanie ciągłości w postaci różniczkowej

$$\boxed{\frac{\partial \varrho}{\partial t} + \nabla \cdot \boldsymbol{J} = 0} \tag{A/8}$$

Postać całkowa tego równania przedstawia się następująco:

$$-\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_T \varrho \, \mathrm{d}\tau = \oint\limits_S J_n \, \mathrm{d}\sigma \tag{A/9}$$

gdzie J_n – składowa normalna wektora $\boldsymbol{J}$.



A/1

Równanie (A/9) orzeka, że przypadająca na jednostkę czasu zmiana ładunku $Q = \int\limits_T \varrho \, \mathrm{d}\tau$ zawartego w obszarze przestrzennym T zamkniętym powierzchnią S jest równa prądowi $I = \oint\limits_S J_n \, \mathrm{d}\sigma$ wypływającemu z tego obszaru, tzn. przepływającemu przez powierzchnię S (rys. A/1).

DODATEK **B**

FUNKCJE FALOWE ATOMU WODORU

Funkcje te, oznaczane symbolami ψ_{nlm}, dogodnie jest przedstawiać w postaci

$$\psi_{nlm}(r, \vartheta, \varphi) = R_{nl}(r)\,\Theta_{lm}(\vartheta)\exp(jm\varphi) \tag{B/1}$$

gdzie: $R_{nl}(r)$ – noszą nazwę funkcji *radialnych*, a $\Theta_{lm}(\vartheta)\exp(jm\varphi)$ – funkcji *kątowych*. Kształt funkcji R_{nl} zależy od wartości liczb kwantowych n i l, natomiast funkcji Θ_{lm} od wartości liczb l i m.

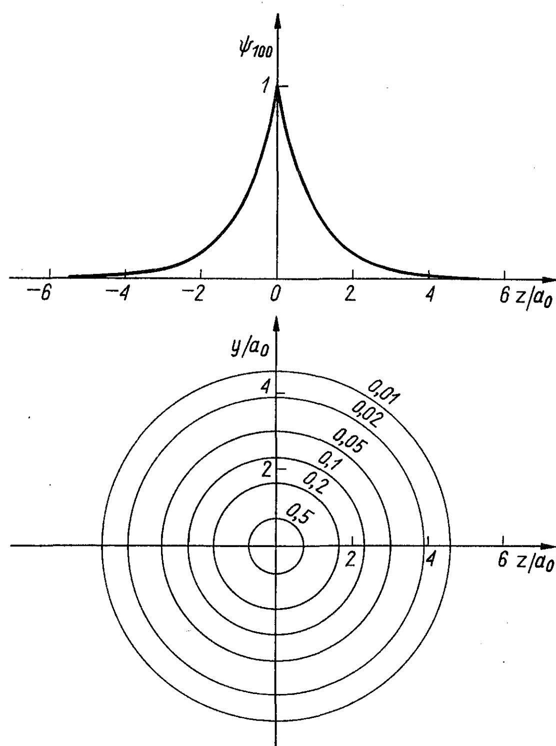
Funkcje te dla $n = 1, 2, 3$ mają następującą postać:

$$\left.\begin{aligned}
R_{10} &= 2\left(\frac{1}{a_0}\right)^{3/2}\exp\left(-\frac{r}{a_0}\right)\\
R_{20} &= \frac{1}{2\sqrt{2}}\left(\frac{1}{a_0}\right)^{3/2}\left(2-\frac{r}{a_0}\right)\exp\left(-\frac{r}{2a_0}\right)\\
R_{21} &= \frac{1}{2\sqrt{6}}\left(\frac{1}{a_0}\right)^{5/2} r\exp\left(-\frac{r}{2a_0}\right)\\
R_{30} &= \frac{2}{81\sqrt{3}}\left(\frac{1}{a_0}\right)^{3/2}\left[27-18\frac{r}{a_0}+2\left(\frac{r}{a_0}\right)^2\right]\exp\left(-\frac{r}{3a_0}\right)\\
R_{31} &= \frac{4}{81\sqrt{6}}\left(\frac{1}{a_0}\right)^{5/2}\left(6-\frac{r}{a_0}\right) r\exp\left(-\frac{r}{3a_0}\right)\\
R_{32} &= \frac{4}{81\sqrt{30}}\left(\frac{1}{a_0}\right)^{7/2} r^2\exp\left(-\frac{r}{3a_0}\right)
\end{aligned}\right\} \tag{B/2}$$

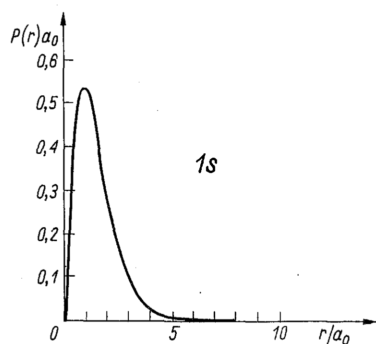
$$\left.\begin{aligned}
\Theta_{00} &= \frac{1}{2\sqrt{\pi}}\\
\Theta_{01} &= \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\vartheta\\
\Theta_{11} &= \frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\vartheta\\
\Theta_{20} &= \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\vartheta-1)\\
\Theta_{21} &= \frac{1}{2}\sqrt{\frac{15}{2\pi}}\sin\vartheta\cos\vartheta\\
\Theta_{22} &= \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\vartheta
\end{aligned}\right\} \qquad \text{(B/3)}$$

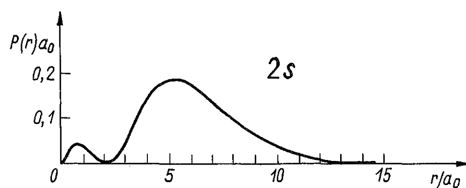
Ze wzorów (B/3) wynika, że funkcje falowe stanów s są sferycznie symetryczne, pozostałe zaś mają własności kierunkowe, tzn. są zależne od kąta ϑ.

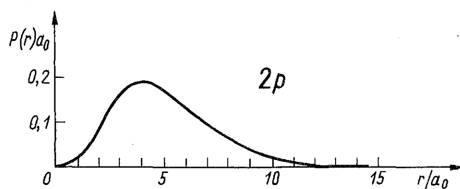
Kształt orbitali $1s$ oraz $2p_0$ pokazano na rys. B/1 i B/3. Wartości funkcji falowych odłożono tu w skali względnej w stosunku do wartości maksymalnej każdej funkcji. W górnej części każdego rysunku przedstawiono zależność funkcji

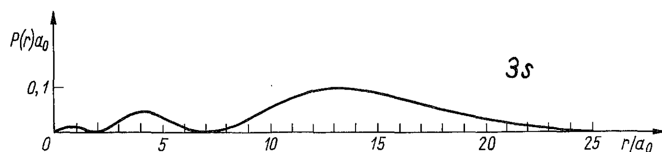


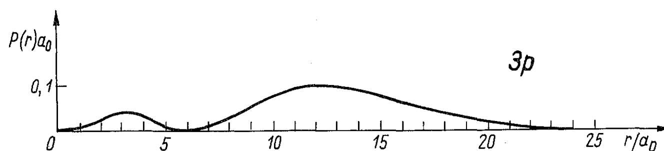
B/1 Kształt orbitalu 1s w atomie wodoru

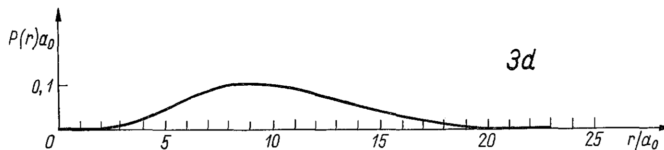




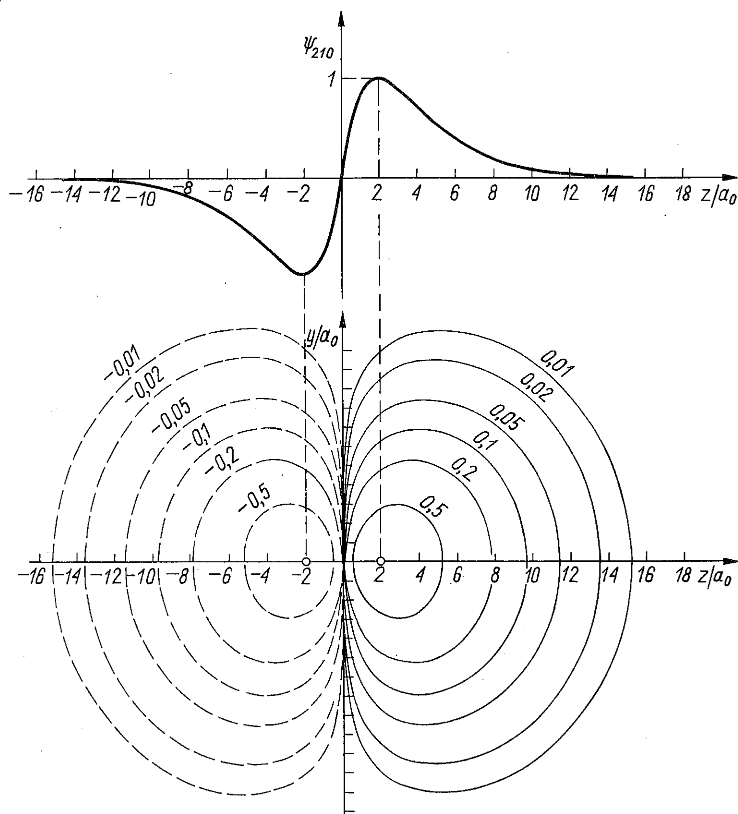








B/2 Rozkład radialnej gęstości prawdopodobieństwa znalezienia elektronu w atomie wodoru w najniższych stanach kwantowych



B/3 Kształt orbitalu $2p_0$ w atomie wodoru

falowej od współrzędnej z, w dolnej zaś – pokazano linie stałych wartości funkcji ψ w płaszczyźnie yz.

Na rysunku B/2 pokazano radialną gęstość prawdopodobieństwa znalezienia elektronu w atomie wodoru w stanach kwantowych o liczbie $n = 1, 2$ oraz 3.

DODATEK **C**

UKŁAD OKRESOWY PIERWIASTKÓW

Schemat zapełniania się kolejnych podpowłok elektronowych przedstawiono w tabl. C/1 (patrz wklejka). Dla każdego pierwiastka podano jego liczbę atomową, symbol chemiczny oraz uproszczoną konfigurację elektronową. W kolumnach pionowych są umieszczone pierwiastki mające identyczne lub bardzo podobne konfiguracje zewnętrznych podpowłok elektronowych. Tworzą one tzw. *grupy* (wzgl. rodziny) układu okresowego, noszące następujące nazwy:

Rodziny główne:

- Ia – wodór i metale alkaliczne (litowce)
- IIa – metale ziem alkalicznych (wapniowce)
- IIIb – glinowce
- IVb – węglowce
- Vb – azotowce
- VIb – siarkowce
- VIIb – halogeny (chlorowce)
- 0 – gazy szlachetne (helowce)

Rodziny dodatkowe (poboczne):

- IIIa – skandowce
- IVa – tytanowce
- Va – wanadowce
- VIa – chromowce
- VIIa – manganowce
- VIII – żelazowce (Fe, Ni, Co) i platynowce (pozostałe 6 pierwiastków)
- Ib – miedziowce
- IIb – cynkowce

Układ okresowy pierwiastków przedstawia się najczęściej w postaci pokazanej w tabl. C/2.

c/1

Okres	Zapełniająca się podpowłoka	Ia	IIa	IIIb	IVb	Vb	VIb	VIIb	0	IIIa	IVa	Va	VIa	VIIa	VIII			Ib	IIb
I	1s	1 H $1s^1$							2 He $1s^2$										
II	2s	3 Li $2s^1$	4 Be $2s^2$																
	2p			5 B $2s^22p^1$	C $2s^22p^2$	7 N $2s^22p^3$	8 O $2s^22p^4$	9 F $2s^22p^5$	10 Ne $2s^22p^6$										
III	3s	11 Na $3s^1$	12 Mg $3s^2$																
	3p			13 Al $3s^23p^1$	14 Cl $3s^23p^2$	15 P $3s^23p^3$	16 S $3s^23p^4$	17 Cl $3s^23p^5$	18 Ar $3s^23p^6$										
IV	4s	19 K $4s^1$	20 Ca $4s^2$																
	3d									21 Sc $3d^14s^2$	22 Ti $3d^24s^2$	23 V $3d^24s^2$	24 Cr $3d^54s^1$	25 Mn $3d^54s^2$	26 Fe $3d^64s^2$	27 Co $3d^74s^2$	28 Ni $3d^84s^2$	29 Cu $4s^1$	30 Zn $4s^2$
	4p			31 Ga $4s^24p^1$	32 Ge $4s^24p^2$	33 As $4s^24p^3$	34 Se $4s^24p^4$	35 Br $4s^24p^5$	36 Kr $4s^24p^6$										
V	5s	37 Rb $5s^1$	38 Sr $5s^2$																
	4d									39 Y $4d^15s^2$	40 Zr $4d^25s^2$	41 Nb $4d^45s^1$	42 Mo $4d^55s^1$	43 Tc $4d^55s^2$	44 Ru $4d^75s^1$	45 Rh $4d^85s^1$	46 Pd $4d^{10}$	47 Ag $5s^1$	48 Cd $5s^2$
	5p			49 In $5s^25p^1$	50 Sn $5s^25p^2$	51 Sb $5s^25p^3$	52 Te $5s^25p^4$	53 J $5s^15p^5$	54 Xe $5s^15p^6$										
VI	6s	55 Cs $6s^1$	56 Ba $6s^2$																
	4f,5d									57–71 Lanta-nowce	72 Hf $5d^26s^2$	73 Ta $5d^36s^2$	74 W $5d^46s^2$	75 Re $5d^56s^2$	76 Os $5d^66s^2$	77 Ir $5d^76s^2$	78 Pt $5d^96s^1$	79 Au $6s^1$	80 Hg $6s^2$
	6p			81 Tl $6s^26p^1$	82 Pb $6s^26p^2$	83 Bi $6s^26p^3$	84 Po $6s^26p^4$	85 At $6s^26p^5$	86 Rn $6s^26p^6$										
VII	7s	87 Fr $7s^1$	88 Ra $7s^2$																
	5f,6d									89–103 Akty-nowce	104 Ku $6d^27s^2$								

Układ okresowy pierwiastków

c/2

Ia	IIa	IIIa		Va	VIa	VIIa	VIII			Ib	IIb	IIIb	IVb	Vb	VIb	VIIb	0
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 A
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 J	54 Xe
55 Cs	56 Ba	57–71 Lantanowce	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89–103 Aktynowce	104 Ku														

Lantanowce	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
Aktynowce	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lw

DODATEK **D**

SZEROKOŚĆ PRZERWY ENERGETYCZNEJ PÓŁPRZEWODNIKÓW

Typ związku	Materiał	Struktura krystaliczna	Przerwa energetyczna	W_g [eV] $T = 300$ K
Pierwiastki	C	diamentu	skośna	5,48
	Si	diamentu	skośna	1,12
	Ge	diamentu	skośna	0,66
	Se	heksagonalna	prosta	2,05
	Te	heksagonalną	prosta	0,34
IV-IV	SiC	blendy cynkowej	skośna	2,20
III-V	AlP	blendy cynkowej	skośna	2,45
	AlAs	blendy cynkowej	skośna	2,15
	AlSb	blendy cynkowej	skośna	1,63
	GaN	wurcytu	prosta	3,44
	GaP	blendy cynkowej	skośna	2,27
	GaAs	blendy cynkowej	prosta	1,43
	GaSb	blendy cynkowej	prosta	0,70
	InN	wurcytu	prosta	2,07
	InP	blendy cynkowej	prosta	1,34
	InAs	blendy cynkowej	prosta	0,36
	InSb	blendy cynkowej	prosta	0,18
II-VI	ZnS	blendy cynkowej	prosta	3,56
	ZnSe	blendy cynkowej	prosta	2,70
	ZnTe	blendy cynkowej	prosta	2,35
	CdS	wurcytu	prosta	2,50
	CdSe	wurcytu	prosta	1,75
	CdTe	blendy cynkowej	prosta	1,45
	HgSe	blendy cynkowej	odwrócona	−0,061
	HgTe	blendy cynkowej	odwrócona	−0,16
IV-VI	PbS	soli kuchennej	prosta	0,40
	PbSe	soli kuchennej	prosta	0,25
	PbTe	soli kuchennej	prosta	0,31

DODATEK E

STOPIEŃ JONIZACJI DOMIESZEK

Prawdopodobieństwo obsadzenia przez elektron poziomu o energii W wyraża funkcja rozkładu Fermiego-Diraca o postaci (8/1). Dotyczy ona przypadku, w którym maksymalna liczba elektronów, które mogą obsadzić dany poziom energetyczny, jest równa liczbie przyporządkowanych temu poziomowi stanów kwantowych. Z taką sytuacją mieliśmy do czynienia przy rozpatrywaniu pasm energetycznych w półprzewodniku. Inaczej jednak wygląda sprawa w przypadku poziomów domieszkowych. Każdemu donorowi odpowiadają na przykład 2 stany kwantowe, różniące się orientacją spinu elektronu. Z tych dwóch stanów tylko jeden może być jednak obsadzony przez elektron, przy czym oba stany mają jednakowe prawdopodobieństwo obsadzenia. Z chwilą obsadzenia danego stanu donor staje się atomem elektrycznie obojętnym i nie jest zdolny do przyjęcia następnego elektronu. O takim poziomie energetycznym mówimy, że jest zdegenerowany, przy czym *współczynnik degeneracji* g jest w rozpatrywanym przypadku równy 2.

Rozważając stopień obsadzenia poziomów zdegenerowanych, należy posługiwać się tzw. *zmodyfikowanym* rozkładem Fermiego-Diraca o postaci

$$f(W) = \frac{1}{1+\frac{1}{g}\exp\frac{W-W_F}{kT}} \tag{E/1}$$

Jeżeli N_d oznacza koncentrację wszystkich donorów, a N_d^+ – koncentrację donorów zjonizowanych (tzn. nie obsadzonych przez elektrony), to koncentracja stanów donorowych obsadzonych przez elektrony wynosi $N_d-N_d^+$. Korzystając ze wzoru (E/1) i podstawiając $W = W_d$ oraz $g = 2$, możemy napisać

$$N_d-N_d^+ = \frac{N_d}{1+\frac{1}{2}\exp\frac{W_d-W_F}{kT}} \tag{E/2}$$

co prowadzi do następującego wzoru wyrażającego stopień jonizacji donorów:

$$\frac{N_d^+}{N_d} = \frac{1}{1+2\exp\left(-\frac{W_d - W_F}{kT}\right)} \tag{E/3}$$

Dla akceptorów współczynnik degeneracji wynosi 4, mamy tu bowiem do czynienia z degeneracją pasma walencyjnego (pasma dziur lekkich i dziur ciężkich) oraz niezależną od tego degeneracją spinową. Koncentracja stanów akceptorowych obsadzonych przez elektrony wynosi N_a^-, a zatem na podstawie wzoru (E/1)

$$\frac{N_a^-}{N_a} = \frac{1}{1+\frac{1}{4}\exp\frac{W_a - W_F}{kT}} \tag{E/4}$$

Przy niezbyt niskich temperaturach (np. przy $T > 250$ K) i niezbyt wielkich koncentracjach domieszek (np. dla krzemu przy koncentracjach nie przekraczających 10^{17} cm^{-3}) wyrażenia w mianownikach wzorów (E/3) i (E/4) są tak bliskie jedności, że można przyjmować, iż wszystkie domieszki są zjonizowane, tzn.

$$N_d^+ = N_d \quad \text{oraz} \quad N_a^- = N_a$$

Przykładowo, przy $T = 300$ K i $N = 10^{17}$ cm^{-3}, otrzymuje się dla krzemu domieszkowanego fosforem

$$N_d^+ = 0{,}967\, N_d$$

a dla krzemu domieszkowanego borem

$$N_a^- = 0{,}993\, N_a$$

DODATEK **F**

QUASI-POZIOMY FERMIEGO

Statystyka Fermiego-Diraca i związane z nią pojęcie poziomu Fermiego odnoszą się do stanu równowagi termodynamicznej. W tych warunkach koncentracje nośników ładunku w półprzewodniku są wyrażone wzorami (8/35) i (8/36). Ze wzorów tych wynika, że zarówno n_0, jak i p_0 zależą od jednego parametru — energii Fermiego W_F.

W stanie nierównowagowym można, formalnie biorąc, wyrazić koncentracje elektronów i dziur za pomocą wzorów analogicznych do (8/35) i (8/36), z tym jednak zastrzeżeniem, że zamiast poziomu Fermiego należy wprowadzić 2 niezależne parametry, z których jeden dotyczy elektronów, drugi zaś dziur. Wielkości te nazywamy *quasi-poziomami Fermiego* i oznaczamy odpowiednio symbolami W_{Fn} i W_{Fp}.

Przez analogię do wyrażeń (8/35) i (8/36) można więc napisać:

$$n = n_0 + \Delta n = n_i \exp \frac{W_{Fn} - W_i}{kT} \qquad \text{(F/1)}$$

$$p = p_0 + \Delta p = n_i \exp \frac{W_i - W_{Fp}}{kT} \qquad \text{(2F/)}$$

Równania te można traktować jako definicje wielkości W_{Fn} i W_{Fp}. W warunkach równowagi termodynamicznej $\Delta n = \Delta p = 0$, a zatem $W_{Fn} = W_{Fp} = W_F$. W warunkach nierównowagowych natomiast W_{Fn} jest na ogół różne od W_{Fp}, przy czym różnica tych quasi-poziomów jest tym większa, im jest większa koncentracja nośników nadmiarowych. Wprowadzenie nośników nadmiarowych powoduje więc rozszczepienie poziomu Fermiego na 2 quasi-poziomy.

W celu wyznaczenia położenia tych quasi-poziomów w stosunku do równowagowego poziomu Fermiego, oprzemy się na równaniach (F/1) i (F/2) oraz

(8/35) i (8/36). Po pewnych przekształceniach otrzymamy

$$W_{Fn} - W_F = kT \ln\left(1+\frac{\Delta n}{n_0}\right) \tag{F/3}$$

$$W_F - W_{Fp} = kT \ln\left(1+\frac{\Delta p}{p_0}\right) \tag{F/4}$$

Jeżeli przykładowo rozpatrujemy półprzewodnik typu p i założymy, że $\Delta n \approx \Delta p > 0$, to jak wynika z podanych wyżej wzorów, quasi-poziom dla elektronów przesunie się w górę w stosunku do równowagowego poziomu Fermiego, natomiast quasi-poziom dla dziur przesunie się w dół. Przy małym poziomie zakłócenia przesunięcie tego ostatniego jest znikomo małe, przesunięcie quasi-poziomu W_{Fn} ma natomiast zasadnicze znaczenie. Wprowadzenie nośników nadmiarowych wpływa więc przede wszystkim na położenie quasi-poziomu dla nośników mniejszościowych. Położenie quasi-poziomu dla nośników większościowych ulega zmianie dopiero przy dużym poziomie zakłócenia.

Z zależności (F/1) i (F/2) wynikają dla półprzewodnika niezdegenerowanego następujące związki między gęstościami prądów J_n i J_p, a odpowiednimi quasi-poziomami W_{Fn} i W_{Fp}:

$$J_n = n\mu_n \nabla W_{Fn} \tag{F/5}$$

$$J_p = p\mu_p \nabla W_{Fp} \tag{F/6}$$

Słuszność tych zależności wynika z następujących przekształceń, które przeprowadzimy przykładowo dla elektronów, ograniczając się przy tym do przepływu jednowymiarowego. Podstawiając wyrażenie (10/15) do wzoru (10/66), otrzymujemy

$$J_n = q\mu_n\left(\frac{n}{q}\,\frac{dW_i}{dx} + U_T\,\frac{dn}{dx}\right) \tag{F/7}$$

Jednocześnie z różniczkowania wyrażenia (F/1) wynika

$$\frac{dn}{dx} = \frac{n}{kT}\left(\frac{dW_{Fn}}{dx} - \frac{dW_i}{dx}\right) \tag{F/8}$$

co po podstawieniu do (F/7) prowadzi do wzoru

$$J_n = n\,\mu_n \frac{dW_{Fn}}{dx} \tag{F/9}$$

będącego jednowymiarowym odpowiednikiem zależności (F/5).

DODATEK 

DANE MATERIAŁOWE NAJWAŻNIEJSZYCH PÓŁPRZEWODNIKÓW

Parametr	Jednostka	T [K]	Si	Ge	GaAs
Stała sieci	nm	300	0,543	0,566	0,565
Koncentracja atomów	cm^{-3}	300	$5{,}00 \cdot 10^{22}$	$4{,}42 \cdot 10^{22}$	$4{,}42 \cdot 10^{22}$
Przerwa energetyczna	eV	0	1,17	0,74	1,52
		300	1,12	0,66	1,43
Masa efektywna elektronów[1]					
podłużna	–	300	0,92	1,64	0,067
poprzeczna	–	300	0,19	0,082	
Masa efektywna dziur[1]					
lekkich	–	300	0,15	0,044	0,068
ciężkich	–	300	0,54	0,28	0,50
Masa efektywna gęstości stanów[1]					
m_c^*	–	300	1,18	0,89	0,067
m_v^*	–	300	0,81	0,29	0,52
Ruchliwość					
elektronów	$cm^2/V \cdot s$	300	1420	3900	8500
dziur	$cm^2/V \cdot s$	300	450	1900	400
Przenikalność elektryczna względna	–	300	11,9	16,0	12,9
Czas życia nośników nadmiarowych	s	300	$2{,}5 \cdot 10^{-3}$	10^{-3}	10^{-8}

[1] Masy efektywne nośników podano w jednostkach względnych (w stosunku do masy spoczynkowej elektronu m_0).

DODATEK 

UZUPEŁNIAJĄCE ZADANIA PROBLEMOWE

Do rozdziału 1

1. Wyprowadź wzór liczbowy uzależniający energię fotonu wyrażoną w eV od długości fali promieniowania wyrażonej w nm. Wykonaj odpowiedni wykres, stosując na obu osiach skalę logarytmiczną, i zaznacz na nim zakres widzialny (obejmuje on fale o długości od 380 do 760 nm) oraz promieniowanie podczerwone i nadfioletowe.

2. Żarówka o mocy 60 W wysyła promieniowanie równomiernie we wszystkich kierunkach. Wyznacz liczbę fotonów padających w ciągu 1 s na znajdującą się w odległości 100 m płaszczyznę o powierzchni 1 cm^2 ustawioną prostopadle do kierunku rozchodzenia się promieniowania. Dla uproszczenia przyjmij, że żarówka emituje światło monochromatyczne o barwie żółtej, o danych jak w probl. 1.3.

3. Tak zwana stała słoneczna jest to strumień energii promieniowania słonecznego padającego na płaszczyznę o powierzchni 1 m^2, ustawioną prostopadle do kierunku promieni słonecznych. Przy czystym niebie wynosi ona na powierzchni morza 1 kW/m^2. Całkowita liczba fotonów słonecznych, przypadających na jednostkę powierzchni i jednostkę czasu, wynosi w tych warunkach $4{,}8 \cdot 10^{17}$ $cm^{-2}s^{-1}$. Wyznacz średnią energię tych fotonów i odpowiadającą jej długość fali świetlnej.

4. Wykaż, że ze wzorów (1/38) wypływa wzór (1/39). Potraktuj zagadnienie jednowymiarowo.

5. Atom wodoru składa się, jak wiadomo, z dodatnio naładowanego protonu i poruszającego się wokół niego elektronu obdarzonego ładunkiem ujemnym. Opierając się na zasadzie nieoznaczoności wyjaśnij, dlaczego elektron, pomimo działającej na niego siły przyciągającej nie spada na jądro, lecz znajduje się w stosunkowo dużej odległości od niego. Oszacuj tę odległość, tzn. wyznacz orientacyjną wartość promienia atomowego, wie-

dząc, że energia kinetyczna elektronu w stanie podstawowym atomu wodoru wynosi 13,6 eV.

6. Z przyjętego przez Plancka postulatu, wyrażonego wzorem (1/31), wynika, że dla $n = 0$ energia oscylatora harmonicznego jest równa zeru. Wyjaśnij, dlaczego założenie to jest niezgodne z poznanymi prawami mechaniki kwantowej. (Poprawną postać wyrażenia na energię oscylatora harmonicznego przedstawia wzór (1/32)).

7. Na podstawie pomiaru wiemy, że prędkość pocisku ($m = 50$ g) i prędkość elektronu są takie same, równe 300 m/s z dokładnością 0,01%. Z jaką maksymalną dokładnością możemy wyznaczyć położenie każdego z tych obiektów, jeżeli położenie jest mierzone jednocześnie z prędkością w tym samym doświadczeniu?

Do rozdziału 3

1. Znając rozkład radialnej gęstości chmury elektronowej $P(r)$ dla stanu podstawowego atomu wodoru (probl. 3.4), oblicz jaka część tej chmury znajduje się w otaczającej proton kuli o promieniu r, a następnie podstaw $r = a_0$, $2a_0$, $3a_0$ i $10a_0$.

2. Na rysunku 3/1 przedstawiono tzw. serię Balmera widma emisyjnego atomu wodoru. Serię taką otrzymuje się przy przejściach elektronu z różnych poziomów energetycznych na jeden określony poziom. Zidentyfikuj ten poziom i sprawdź poprawność podanych na rysunku długości fali. Wykonaj analogiczny rysunek dla innej serii widmowej.

Uwaga: Niewielka niedokładność, jaka może pojawić się w obliczeniach, wynika z tego, że rozpatrywana w podręczniku teoria atomu wodoru opiera się na założeniu nieskończenie wielkiej masy protonu.

Do rozdziału 5

1. Na podstawie wyników uzyskanych w probl. 5.1 wyznacz gęstość krzemu.

2. Promień rdzenia atomowego krzemu w krysztale wynosi 39 pm. Oblicz, jaki procent objętości kryształu krzemu zajmują rdzenie atomowe. Czy objętość kryształu jest raczej wypełniona, czy też raczej pusta?

Do rozdziału 8

1. Wyznacz koncentracje samoistne krzemu, germanu i arsenku galu w temperaturze 300 i 350 K, wykorzystując dane zestawione w dodatku G. Nie uwzględniaj zależności W_g od temperatury.

2. Wyznacz koncentrację samoistną diamentu w temperaturze pokojowej, przyjmując $m_c = m_v = m_0$. Jak duży musiałby być kryształ diamentu, aby, statystycznie rzecz biorąc, znalazła się w nim 1 para elektron-dziura?

3. W metalu w temperaturze $T = 0$ K w najwyższym pasmie energe... zny... wszystk e poziomy leżące poni... i pozi mu Fermiego są cał-

kowicie obsadzone przez elektrony, natomiast poziomy leżące powyżej W_F są nie obsadzone. W temperaturach różnych od zera podział poziomów na obsadzone i nie obsadzone nie przebiega tak ostro. Istnieje wówczas pewien przedział energii, w którym prawdopodobieństwo obsadzenia zmienia się od wartości bliskiej jedności (wynoszącej np. 0,99) do wartości bliskiej zera (np. 0,01). Wyznacz szerokość tego przedziału w jednostkach kT.

4. W atomie sodu w temperaturze $T = 0$ K obsadzone są całkowicie poziomy 1*s*, 2*s* i 2*p*, natomiast poziom 3*s* jest obsadzony w 50%. Oznacza to, że poziom Fermiego pokrywa się z poziomem 3*s*. Wyznacz prawdopodobieństwa obsadzenia poziomów 2*p* i 3*p* w temperaturze pokojowej. Wykorzystaj dane z rys. 3/8 i 6/2.

5. Dla krzemu o koncentracji donorów $N_d = 10^{15}$ cm^{-3} w temperaturze pokojowej wyznacz koncentracje równowagowe n_0 i p_0 oraz położenie poziomu Fermiego dla $T = 300$ i 350 K. Narysuj obok siebie modele pasmowe tego półprzewodnika dla obu wartości temperatury. Wyjaśnij dlaczego ze wzrostem temperatury poziom Fermiego przesuwa się w kierunku poziomu samoistnego.

6. Dla krzemu o koncentracji domieszek $N_d = 10^{16}$ cm^{-3} i $N_a = 10^{15}$ cm^{-3} wyznacz koncentracje równowagowe n_0 i p_0 oraz położenie poziomu Fermiego dla $T = 300$ K. W związku z tym, że n_0 jest mniejsze od N_d, wyjaśnij co się stało z pozostałymi elektronami z poziomów donorowych. Pokaż to na modelu pasmowym.

7. Dla półprzewodnika skompensowanego o koncentracji domieszek $N_d = N_a = 10^{16}$ cm^{-3} wyznacz koncentracje równowagowe n_0 i p_0 oraz położenie poziomu Fermiego dla $T = 300$ K. Narysuj model pasmowy tego półprzewodnika i zaznacz na nim, które poziomy energetyczne są obsadzone przez elektrony i jaki jest stopień ich obsadzenia.

ROZWIĄZANIA PROBLEMÓW

1.3. $n = 10$. Biorąc pod uwagę fakt, że część światła ulega odbiciu od poszczególnych warstw gałki ocznej, a znaczną część zostaje pochłonięta w gałce zanim dotrze do siatkówki, dochodzimy do wniosku, że prawie każdy foton zaabsorbowany przez receptor wzrokowy powoduje jego pobudzenie.

1.4. $\lambda = 1{,}17 \cdot 10^{-24}/\sqrt{m\, W_k}$; $\lambda = 5{,}5$ pm.

1.5. $d = 91$ pm

1.6. $D = d + \frac{L\lambda}{\pi d}$, gdzie λ – długość fali elektronowej. Szerokość plamki staje się najmniejsza przy $d = d_0 = \sqrt{L\lambda/\pi}$; $D_{min} = 2d_0$. W rozpatrywanym przypadku $d_0 = 2$ µm.

1.7. $\frac{\Delta W}{W} \geqslant 2 \cdot 10^{-8}$. Jednostka długości jest więc zdefiniowana z bardzo dużą dokładnością.

2.4. $x = 0{,}64$ nm; $\lambda = 1{,}74$ nm, a zatem gęstość prawdopodobieństwa znalezienia elektronu w obszarze klasycznie zabronionym spada do wartości równej 1% wartości początkowej na drodze o długości mniejszej niż połowa długości fali elektronowej w obszarze klasycznie dozwolonym.

2.5. $B = \frac{k_1 - k_2}{k_1 + k_2} A$; $C = \frac{2k_1}{k_1 + k_2} A$.

2.7. Dla $k_2 d > 2$, $T \approx \frac{4}{\eta^2} \exp(-2k_2 d)$.

2.8. $A_n = \sqrt{2/a}$.

2.9. $k_2^2 + 2k_1 k_2 \operatorname{ctg} k_1 a - k_1^2 = 0$, gdzie $k_1 = \sqrt{\frac{2m}{\hbar^2} W}$, $k_2 = \sqrt{\frac{2m}{\hbar^2}(U_0 - W)}$.
Gdy $U_0 \to 0$, wówczas $k_2 \to \infty$, co prowadzi do wniosku, że $\operatorname{ctg} k_1 a \to \infty$, czyli że $k_1 a = n\pi$.

3.1. $R = \frac{q^2}{8\pi hc\,\varepsilon_0\,a_0}$; $U_j = Rhc = 13{,}6\ \text{eV}$.

3.3. $A^2 = \frac{1}{\pi}\left(\frac{1}{a_0}\right)^3$.

3.4. Najbardziej prawdopodobna odległość elektronu od jądra wynosi a_0; $\langle r \rangle = \frac{3}{2} a_0$.

4.1. Sieć skośna ($a \neq b$; $\gamma \neq 90°$), sieć prostokątna ($a \neq b$; $\gamma = 90°$), sieć kwadratowa ($a = b$; $\gamma = 90°$), sieć heksagonalna ($a = b$; $\gamma = 60°$), sieć prostokątna centrowana.

4.2. $N = 2{,}65 \cdot 10^{22}\ \text{cm}^{-3}$; $\varrho = 1{,}02\ \text{g/cm}^3$.

5.1. $d = 235\ \text{pm}$; $N = 5 \cdot 10^{22}\ \text{cm}^{-3}$.

6.1. $\lambda < \frac{hc}{W_g} = \frac{1240}{W_g}$ [nm], przy czym w ostatnim wzorze szerokość przerwy energetycznej W_g jest wyrażona w elektronowoltach. Półprzewodniki o $W_g < 1{,}63\ \text{eV}$ są nieprzezroczyste, o $W_g > 3{,}26\ \text{eV}$ – przezroczyste, a o W_g zawartym w powyższych granicach częściowo przezroczyste.

7.1. $p = 7{,}6 \cdot 10^{-28}\ \text{kg} \cdot \text{m/s}$; pęd krystaliczny na granicy strefy Brillonina jest rzędu $10^{-24}\ \text{kg} \cdot \text{m/s}$.

7.2. $N \approx 5 \cdot 10^{17}\ \text{cm}^{-3}$.

8.4. $\frac{1}{n_i}\frac{dn_i}{dT} = \frac{3}{2T} + \frac{W_{g0}}{2kT^2}$; dla krzemu w temperaturze pokojowej $\frac{1}{n_i}\frac{dn_i}{dT} = 0{,}08\ \text{K}^{-1}$.

8.5. $N_d = 4{,}5 \cdot 10^{18}\ \text{cm}^{-3}$.

9.3. $1 - f(W_t) = \frac{n_i + \Delta p}{n_0 + 2\Delta p}$; gdy $\Delta p \to 0$, wówczas $1 - f(W_t) \to \frac{n_i}{n_0}$; gdy natomiast $\Delta p \to \infty$, wówczas $1 - f(W_t) \to \frac{1}{2}$.

10.1. $\sigma_i = qn_i(\mu_n + \mu_p)$; dla krzemu w temperaturze pokojowej $\sigma_i = 3\ (\mu\Omega \cdot \text{cm})^{-1}$; $\varrho_i = 330\ \text{k}\Omega \cdot \text{cm}$; $\frac{1}{\sigma_i}\frac{d\sigma_i}{dT} = 0{,}072\ \text{K}^{-1}$.

10.2. $\mu = 37\ \text{cm}^2/\text{V} \cdot \text{s}$; $\tau = 2{,}7 \cdot 10^3\ \text{s} = 45\ \text{min}$; $E = 1\ \text{mV/cm}$.

10.4. $\sigma = 2{,}9\ (\Omega \cdot \text{cm})^{-1}$; $\varrho = 3{,}5\ \Omega\text{cm}$; $\frac{1}{\sigma}\frac{d\sigma}{dT} = -0{,}007\ \text{K}^{-1}$.

12.1. $E = 91\ \text{V/cm}$; $I = 820\ \text{A}$; $\tau = 1{,}56 \cdot 10^{-13}\ \text{s}$.

12.2. $\frac{\Delta E}{E_0} = \frac{(b+1)\Delta n}{(n_0 + \Delta n)\, b + p_0 + \Delta n}$; dla półprzewodnika samoistnego $\frac{\Delta E}{E_0} = \frac{\Delta n}{n_i + \Delta n}$; dla silnie domieszkowanego półprzewodnika typu n przy małym poziomie zakłócenia

$$\frac{\Delta E}{E_0} = \frac{b+1}{b}\frac{\Delta n}{n_0}.$$

12.3. $Q = I\tau_n$

12.5. $\alpha_T = \dfrac{1}{\operatorname{ch}\dfrac{w}{L_n}}$

12.6. $\tau = \dfrac{w^2}{2D_n}$; $Q = I\tau$

12.10.

$\Delta n(0) = \dfrac{SL_n}{D_n}$ Przy małym poziomie zakłócenia $E(x) = \dfrac{U_T(b-1)}{p_0}\left|\dfrac{dn}{dx}\right|$;
$U = \dfrac{U_T(b-1)}{p_0}\Delta n(0)$. Dla zadanych warunków $\Delta n(0) = 9{,}2 \cdot 10^{13}\,\mathrm{cm}^{-3}$;
$U = 4{,}8$ mV.

WYKAZ LITERATURY

LITERATURA W JĘZYKU POLSKIM

1. Blinowska A., Blinowski J., Gorzkowski W.: *Fale, cząstki, atomy.* Wyd. Szkolne i Pedagogiczne, Warszawa 1977.
2. Boncz-Brujewicz W. L., Kałasznikow S. G.: *Fizyka półprzewodników.* PWN. Warszawa 1985.
3. Eisberg R., Resnick R.: *Fizyka kwantowa atomów, cząsteczek, ciał stałych, jąder i cząstek elementarnych.* PWN, Warszawa 1983.
4. Enge H. A., Wehr M. R., Richards J. A.: *Wstęp do fizyki atomowej.* PWN, Warszawa 1983.
5. Figielski T.: *Zjawiska nierównowagowe w półprzewodnikach.* PWN, Warszawa 1980.
6. Ginter J.: *Wstęp do fizyki atomu, cząsteczki i ciała stałego.* PWN, Warszawa 1979.
7. Holden A.: *Ciało stałe.* PWN, Warszawa 1969.
8. Kittel C.: *Wstęp do fizyki ciała stałego* (tłumaczenie wydania amerykańskiego z 1966 r.). PWN, Warszawa 1976.
9. Kołos W.: *Chemia kwantowa.* PWN, Warszawa, 1978.
10. Kołos W.: *Elementy chemii kwantowej sposobem niematematycznym wyłożone.* PWN, Warszawa 1984.
11. Levine S. N.: *Fizyka kwantowa w elektronice.* PWN, Warszawa 1968.
12. Norwood J.: *Fizyka współczesna.* PWN, Warszawa 1982.
13. Penkala T.: *Zarys krystalografii.* PWN, Warszawa 1976.
14. Sukiennicki A., Zagórski A.: *Fizyka ciała stałego.* WNT, Warszawa 1984.
15. Świt A., Pułtorak J.: *Przyrządy półprzewodnikowe.* WNT, Warszawa 1979.
16. Wert C. A., Thomson R. M.: *Fizyka ciała stałego.* PWN, Warszawa 1974.
17. Wichmann E. H.: *Fizyka kwantowa.* PWN, Warszawa 1973.
18. Wilkes P.: *Fizyka ciała stałego dla metaloznawców.* PWN, Warszawa 1979.
19. van der Ziel A.: *Podstawy fizyczne elektroniki ciała stałego.* WNT, Warszawa 1980.

WYBRANE POZYCJE LITERATURY OBCOJĘZYCZNEJ

1. Adler R. B., Smith A. C., Longini R. L.: *Introduction to semiconductor physics.* J. Wiley, New York 1964.
2. Heywang W., Pötzl H. W.: *Bänderstruktur und Stromtransport.* Springer-Verlag, Berlin 1976.
3. Kittel C.: *Introduction to solid state physics.* Ed. 5., J. Wiley, New York 1976.
4. Madelung O.: *Introduction to solid-state theory.* Springer-Verlag, Berlin 1978.

6. Müller R.: *Grundlagen der Helbleiter-Elektronik*. Springer-Verlag, Berlin 1984.
7. Paul R.: *Halbleiterphysik*. VEB Verlag Technik, Berlin 1974.
8. Phillips J. C.: *Bonds and bands in semiconductors*. Academic Press, New York 1973.
9. Seeger K.: *Semiconductor physics*. Springer-Verlag, Wien 1982.
10. Sze S. M.: *Physics of semiconductor devices*. J. Wiley, New York 1981.
11. Wang F. F. Y.: *Introduction to solid state electronics*. North-Holland Publ. Comp., Amsterdam 1980.
12. Wang S.: *Solid state electronics*. McGraw-Hill, New York 1966.
13. Warner R. M. Jr., Grung B. L.: *Transistors. Fundamentals for the integrated-circuit engineer*. J. Wiley, New York 1983.

LITERATURA DOTYCZĄCA DANYCH MATERIAŁOWYCH

Landolt-Börnstein: *Numerical data and functional relationships in science and technology*. Vol. **17**: *Semiconductors*. Springer-Verlag, Berlin.
– Subvolume **a**: *Physics of group IV elements and III–V compounds*. 1982.
– Subvolume b: *Physics of II–VI* and *I–VII compounds, semimagnetic semiconductors*. 1982.
– Subvolume **c**: *Technology of Si, Ge and SiC*. 1984.
– Subvolume **e**: *Physics of non-tetrahedrally bonded elements and binary compounds I*. 1983.
– Subvolume **f**: *Physics of non-tetrahedrally bounded binary compounds II*. 1983.

SKOROWIDZ

Akceptor 170
ambipolarne równanie transportu 282
ambipolarny współczynnik dyfuzji 283
amplitudowa funkcja falowa 44
anizotropia 80
„antysite defects" 151
atom wodoropodobny 111
atomy międzywęzłowe 86
Augera zjawisko 156

Baza krótka 295
– sieci 81
Blocha twierdzenie 123
Boltzmanna równanie 269
boltzmannowskie przybliżenie 178
Bragga warunek 30
Brillouina strefy 132

Całka Fermiego 181
ciało amorficzne 79
– krystaliczne 79
– stałe 79
czas relaksacji dielektrycznej 278
– – zderzeniowej 236
– życia nośników nadmiarowych 205, 216
częstotliwości rezonansowe 61

Defekty sieci krystalicznej 86
degeneracja spinowa 74
Dembera zjawisko 301
długość fali 26
– wiązania 99
dolina boczna 152
– centralna 152
domieszka izoelektronowa 171
– wodoropodobna 164
– zlokalizowana 171
domieszki 163
– głębokie 164, 209
– płytkie 164, 209
domieszkowanie kryształów 88
donor 166
drgania sieci krystalicznej 89
droga dyfuzji 291
doświadczenie Davissona i Germera 35
dualizm falowo-korpuskularny 34
duży poziom zakłócenia 216
dyfrakcja 30
dyfuzja 224, 251
– dwuskładnikowa 299
dyslokacja 86
dyspersja fal 120
dywergencja 258
dziura 137
dziury ciężkie 152
– lekkie 152

Efekt Gunna 247
efektywna koncentracja domieszki 190
– – samoistna 197

Einsteina relacja 254
ekstrakcja nośników 201
elektron 21
– przewodnictwa 109
– swobodny 44
elektronowolt 24
elektrony gorące 246
– walencyjne 76
elektroujemność 93
emisja dziury 210
– elektronu 210
energia Fermiego 176
– jonizacji 65
– wiązania 93

Fala 24
– elektromagnetyczna 29
– harmoniczna 25
– płaska 24
fale de Broglie'a 35
fenomenologiczna teoria transportu 237
Fermiego-Diraca rozkład 176
– energia 176
fermion 73
fluktuacje 175
fonon 90
– akustyczny 90
– optyczny 90
foton 31
funkcja falowa 42
– rozkładu Fermiego-Diraca 176
funkcje Blocha 123

Gaussa prawo 258
gaz elektronowy 109
generacja nośników 199
– pary elektron-dziura 140
gęstość chmury elektronowej 69
– prawdopodobieństwa 43
– stanów 178
gradient 253
Gunna efekt 247

Hybrydyzacja 100

Interferencja 30

Jonizacja atomu domieszkowego 165
– zderzeniowa 249

Kierunek kwantyzacji 71
komórka elementarna 82, 83
koncentracja nośników 174
– równowagowa 178
– samoistna 184
koncentracje nadmiarowe 200
konfiguracja elektronowa atomu 77
kontakt rekombinacyjny 230
krawędź absorpcji 119, 154
kreacja dziury 140
Kroniga-Penney'a model 124
kryształy mieszane 150
kwant energii 31

Laplasjan 259
liczba atomowa 62
– falowa 26
– kwantowa 57, 63
– – główna 70
– – magnetyczna 71
– – orbitalna 71
– spinowa 73
luka węzłowa 86

Ładunek elementarny 21
– przestrzenny 257

Masa efektywna 135
– – gęstości stanów 179
– – podłużna 160
– – poprzeczna 160
mały poziom wstrzykiwania 290
– – zakłócenia 205, 214
Maxwella równanie 302
mechanika kwantowa 41
metale 107, 110
Millera wskaźniki 85
model atomu Bohra 63
– Kroniga-Penney'a 124
modulacja przewodności 234
monokryształ 79

Napięcie termiczne 254
natężenie fali 28
normowanie potencjału 24
nośniki mniejszościowe 189
– nadmiarowe 200
– większościowe 189

Obsadzenie stanów 179
oddziaływanie spinowo-orbitalne 153
odwrócona struktura pasmowa 161
„ogony" pasm 194
opór elektryczny 234
orbital antywiążący 99
– atomowy 71
– molekularny 98
– wiążący 98
osie krystalograficzne 82

Paczka falowa 27
pasmo odszczepione 152
– paraboliczne 136
– – przewodnictwa 116
– sferyczne 148
– walencyjne 116
Pauliego zakaz 73
pęd krystaliczny 129
pierwsza strefa Brillouina 132
płaszczyzna falowa 25
– sieciowa 85
podpowłoka 75
Poissona równanie 258
pole wbudowane 262
– wewnętrzne 262
potencjał efektywny 122
powierzchnia izoenergetyczna 121
powinowactwo elektronowe 94
powłoki elektronowe 74
poziom akceptorowy 170
– domieszkowy 163
– donorowy 166
– Fermiego 176
– próżni 112
– samoistny 187
półprzewodnik 18, 115, 144
– domieszkowy 151, 188
– niezdegenerowany 182
– o przewodnictwie dziurowym 151
– – – elektronowym 151
– samoistny 150, 184
– skompensowany 192
– typu *n* 151, 189
– – *p* 151, 189
półprzewodniki pierwiastkowe 149
praca wyjścia 32
prawo działania mas 183
– Ohma 233
prąd dyfuzyjny 224, 251
prąd elektryczny 223
– pojemnościowy 225
– przesunięcia 224
– przewodzenia 224
– termodyfuzyjny 224
– unoszenia 224, 227
prędkość fazowa 26
– grupowa 28
– nasycenia 229
– rekombinacji powierzchniowej 221
– termiczna 226
– unoszenia 228
profil domieszkowania 261
promień Bohra 64
prosta sieciowa 81
przejście bezpromieniste 65
– optyczne 65
– proste 138, 154
– skośne 138, 159
przepływ jednowymiarowy 232
– stacjonarny 225
przerwa energetyczna 116
– – prosta 151
– – skośna 157
– – ujemna 161
– – zerowa 161
przestrzeń *k* 144
– wektora falowego 144
przewodność elektryczna 234
– samoistna 234
przybliżenie boltzmannowskie 178
– jednoelektronowe 74, 122
pseudopęd 129
pudło potencjału 59
pulsacja fali 26
pułapki 208
pułapkowanie nośników 214

Quasi-cząstki 90
– – pęd 90
– – poziomy Fermiego 313

Radialna gęstość prawdopodobieństwa 70
rdzeń atomowy 76
rekombinacja Augera 218
– bezpośrednia 201
– międzypasmowa 201

rekombinacja niepromienista 155
– nośników 199
– pośrednia 208
– powierzchniowa 219
– promienista 155
– Shockley'a-Reada-Halla 211
relaksacja 199
– dielektryczna 277
– zderzeniowa 236
rozpraszanie międzydolinowe 247
– na fononach 237
– – swobodnych nośnikach 241
– – zjonizowanych domieszkach 237
– wewnątrzdolinowe 247
równania Boltzmanna 269
– ciągłości 273
– Maxwella 302
równanie ruchu 22
– Schrödingera 43
– – niezależne od czasu 44
równowaga termodynamiczna 174
ruchliwość 229
– grupowa 283
– nośników 235
ruchy termiczne 225

Schrödingera równanie 43
sieć centrowana 82
– krystaliczna 80
– punktowa 81
– translacyjna 81
spin 73
stała Plancka 31
– Rydberga 63
– sieci odwrotnej 132
stan niestacjonarny 60
– podstawowy 64
– quasi-neutralności 263
– stacjonarny 44
– związany 56
stany powierzchniowe 219
– wzbudzone 64
statystyka Fermiego-Diraca 176
stechiometryczny skład 149
struktura blendy cynkowej 105
– diamentu 100
– subtelna linii widmowych 72
– wurcytu 107
studnia potencjału 56
symetria translacyjna 81
– tetraedryczna 101
szybkość generacji 201
– rekombinacji 202

Temperatura elektronowa 246
tensor 148
– odwrotności masy efektywnej 148
teoria pasmowa 110
termodyfuzja 224
twierdzenie Blocha 123

Układ okresowy pierwiastków 308
układy krystalograficzne 83

Wakans 86
warstwa pasywująca 220
wartościowość 76
warunek Bragga 30
– neutralności elektrycznej 189, 212
wektor falowy 26
wektory falowe zredukowane 132
– translacji 82
węzły sieci 81
widmo energetyczne 32
wiązanie heteropolarne 93
– homopolarne 93
– jonowe 94
– kowalencyjne 97
– molekularne 93
– niewysycone 220
– spolaryzowane 103
wskaźniki Millera 85
współczynnik degeneracji 311
– dyfuzji 253
– jonizacji zderzeniowej 249
– jonowości 103
– odbicia 53
– rekombinacji 202
– sprawności transportu 296
– transmisji 53
wstrzykiwanie nośników 201, 290
wurcyt 107
wychwyt dziury 210
– elektronu 209
– jądrowy elektronu 69
wypadkowa szybkość procesów rekombinacyjno-generacyjnych 202
– – rekombinacji 202

Zakaz Pauliego 73
zależność dyspersyjna 120

założenie quasi-neutralności 266
zanieczyszczenia 88
zasada korespondencji 41
– nieoznaczoności 38
– nierozróżnialności 176
– odpowiedniości 41
– równowagi szczegółowej 175
zjawisko Augera 156
– fotoelektryczne 32
– tunelowe 53
złącze *p–n* 261
zmodyfikowany rozkład Fermiego-Diraca 311
zwężenie przerwy energetycznej 197